Expert Systems With Applications 230 (2023) 120577

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Sentiment analysis of COVID-19 cases in Greece using Twitter data

Loukas Samaras *, Elena García-Barriocanal [1], Miguel-Angel Sicilia [2]

Computer Science Department, Polytechnic Building, University of Alcalá, Ctra. De Barcelona km. 33.6, 28871 Alcalá de Henares (Madrid), Spain

A R T I C L E I N F O

Keywords:
Sentiment analysis
Twitter
Pandemic
Web data
Public health

A B S T R A C T

Background: Syndromic surveillance with the use of Internet data has been used to track and forecast epidemics for the last two decades, using different sources from social media to search engine records. More recently, studies have addressed how the World Wide Web could be used as a valuable source for analysing the reactions of the public to outbreaks and revealing emotions and sentiment impact from certain events, notably that of pandemics.
Objective: The objective of this research is to evaluate the capability of Twitter messages (*tweets*) in estimating the sentiment impact of COVID-19 cases in Greece in real time as related to cases.
Methods: 153,528 tweets were gathered from 18,730 Twitter users totalling 2,840,024 words for exactly one year and were examined towards two sentimental lexicons: one in English language translated into Greek (using the Vader library) and one in Greek. We then used the specific sentimental ranking included in these lexicons to track i) the positive and negative impact of COVID-19 and ii) six types of sentiments: *Surprise*, *Disgust*, *Anger*, *Happiness*, *Fear* and *Sadness* and iii) the correlations between real cases of COVID-19 and sentiments and correlations between sentiments and the volume of data.
Results: Surprise (25.32%) mainly and secondly *Disgust* (19.88%) were found to be the prevailing sentiments of COVID-19. The correlation coefficient (R^2) for the Vader lexicon is −0.07454 related to cases and −0.,70668 to the tweets, while the other lexicon had 0.167387 and −0.93095 respectively, all measured at significance level of $p < 0.01$. Evidence shows that the sentiment does not correlate with the spread of COVID-19, possibly since the interest in COVID-19 declined after a certain time.

1. Introduction

From 2004, several studies have reported real-time estimation and prediction of epidemics with the use of data from the Internet. That previous research is in most cases directed to computing and analyzing the volume of this data to estimate the development of an infection in real time and to create rules for forecasts a few weeks ahead and, in most cases, a statistically significant correlation was found between the real development of epidemics and the volume of Internet data. It is observed (Samaras et al., 2020) that messages from social media (Twitter, Facebook, etc.) were used in over the half (50.89%) of the current literature from 2004 until 2020 in order to detect and predict the spread of infectious diseases, such as influenza, malaria, measles, etc., while sentiment analysis in Twitter is a scientific research field which has relatively recently attracted an important research interest (Giachanou and Crestani, 2017).

In a different direction, already from 1986, a major concern was expressed by Marvin Minksy (Minsky, 1986): the question is not weather machines can have any emotions, but weather can be intelligent without any emotions. Within this framework, sentiment analysis concerns the collection of a large amount of sentiment data, which can be useful in understanding the opinion of the people about a variety of topics, as these are discussed in the Web (Pantic et al., 2006). Sentiment analysis of social media has been discussed during the last decade (Agarwal et al., 2011) by providing two major methodologies, separately or combined: knowledge-based systems based upon linguistics rules and statistical machine algorithms.

From the year 2017, research has been exceeded on the fields of the relationship between emotion models and the artificial Intelligence (Bartneck et al., 2008), as well as the multilingual sentiment analysis (Lo et al., 2017). Furthermore, during the last couple of years, sentiment analysis of the Coronavirus disease of 2019 (COVID-19) pandemic has

* Corresponding author.
E-mail addresses: lsamaras@ath.forthnet.gr (L. Samaras), garciab@uah.es (E. García-Barriocanal), msicilia@uah.es (M.-A. Sicilia).
[1] http://orcid.org/0000-0001-6752-9599.
[2] http://orcid.org/0000-0003-3067-4180.

https://doi.org/10.1016/j.eswa.2023.120577
Received 13 February 2023; Received in revised form 29 April 2023; Accepted 27 May 2023
Available online 7 June 2023

raised renewed interest with an effort to gauge the performance of social media for sentiment metrics in association with the coronavirus pandemic, such as Twitter (Agarwal et al., 2011; Bartneck et al., 2008; Lo et al., 2017; Imran et al., 2022; Chen et al., 2022a,b; Huang et al., 2022; Arbane et al., 2023; Kruspe et al., 2022; Okango and Mwambi, 2022; Boon-Itt and Skunkan, 2020; Ridhwan and Hargreaves, 2021; Garcia and Berton, 2021) or Facebook (Catapang and Cleofas, 2022). Recent research also showed (Arbane et al., 2023) that 71% of Twitter users say they use it as source of news, which is evidence in favor of the argument that the effect of Twitter on shaping public opinion is significant.

Apart from the above, many researchers suggest the use of modern neural networks and especially Long Short-Term Memory (LSTM) networks to track and evaluate sentiments or other real-life applications of soft computing techniques in different research fields, such as the following: for stock price prediction (Swathi et al., 2022), e-commerce (Yadav et al., 2023), social media by using fusion model based on attention mechanism (Hongjie et al., 2022), aquacultural engineering (Ashkan et al., 2021), weather and climate with the application of spatiotemporal models for nonlinear distributed thermal processes (Yajun et al., 2020) or with the forecast of rainfall distribution based on fixed sliding window (Chen et al., 2022a,b).

While most of the analysis and discussion pertains the English language or more widely-spoken languages across the world, previous research has not been much applied to other languages, such as Greek (Tsakalidis et al., 2018; Kydroset al., 2021; Kalamatianos et al., 2015). Greek language is not spoken all over the world, but it is the richest language of the world in terms of vocabulary (Github, 2021) and structure, consisting of more than 575,000 words, 7.32% more than the second (Czech) and 22.50% more than the third one (Hebrew/Israel).

In general, sentiment analysis uses a lexicon to identify polarity in text fragments. The sentiment lexicon is usually a set of tokens including the key terms with a specific affective state, either positive, negative or neutral. These distinctions unfold the results of the procedure of classifying a piece of text with respect to its sentiment (He & Zheng, 2019). Most popular (Gitub, cjhutto vaderSentiment, 2022) sentimental lexicons include a small portion of the entire vocabulary of each language. For example, the Vader Sentiment analysis tool ("Valence Aware Dictionary and sEntiment Reasoner", Github, 2020) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media (Github, 2022). This tool is written in Python programming language for use in Twitter and its English lexicon consists of 7,520 words. Another lexicon created by Adam Tsakalidis in collaboration with Symeon Papadopoulos with the contribution of Ourania Voskaki, Kyriaki Ioannidou and Christina Boididou (Tsakalidis et al., 2018; Github, 2020) and used by Nick Krystallis (Github, 2022) consists of 2,316 words and was used by programming R-language algorithms (Orduz, 2018).

The aforesaid lexicons mostly include individual words, without suffixes or prefixes except from the Greek lexicon used by N. Krystallis, which contains 287 suffixes and prefixes. This means that this lexicon can track<3,000 words of the total Greek vocabulary. On the other hand, the English Vader lexicon, translated into Greek provides no suffixes and prefixes at all. Unlike the English language (and possibly other languages), the Greek language has a vocabulary consisting of many words that have different spelling in terms of the various grammar types, such as the verb, person, tense, grammatical number, participles, nouns, adjectives, etc. Particularly, the words "I love", you love, they love" in English has the same word "love" after the personal pronouns, which denote the person of speaking or writing. On the contrast, the Greek language does not have this structure and different words are used to describe the above, e.g., "αγαπώ, αγαπάς, αγαπούν" accordingly. This mean that, to find the person, we must know the suffix of the word: "-ώ, -άς, -ούν", and if we don't, we can only track the Greek term only in the first-person, e.g., only the English equivalent of "I love" and not the others, such as "you love", "he loves", etc. This way, the exact person of the word meaning cannot be tracked in the Greek language, as the verb "love" is the same word for all grammatical persons in English, but different for Greek. The same can be said for the other grammatical variations of words, such as grammatical numbers, tenses, voices, cases and grammatical moods.

With our work we examine the detection of sentiments regarding the COVID-19 in Greece by using Twitter messages written in the Greek language and try to deal with two specific problems:

The structure of the Greek language, the grammatical types of which is different from more popular and more often-spoken languages, such as the English language, as described above.

The decline of the interest over time regarding the COVID-19 pandemic, as this was observed, not only through Twitter messages, but also through Google Trends. This resulted in a significantly less amount of Twitter messages after a certain period and, consequently, in less sentiments.

To address the abovementioned problems, the main goals of our study are not to predict the trend of the mood among people at a certain time or case, but to showcase a set of methods of unveiling the sentiments and their relation with the impact of the recent pandemic in Greece, accounting for the following aspects:

Including both knowledge-based techniques based on the establishment of linguistic rules and techniques based on statistical measurements.

Comparing this impact based on a less-spoken language and the sentiment ranking from different sentiment models, concretely, we used the lexicon of Vader and of Nick Krystallis (Nkryst lexicon).

Extending the lexicons to have the capacity of tracking more words with person identification of the total Greek vocabulary inside the messages based on the term morpheme (μόρφημα) or theme (θέμα) of the morphology and linguistic theory.

Examining a long period of an entire year, since in the current literature most researches have been conducted with data of smaller periods.

Regarding the above-mentioned models, it must be said that both lexicons are capable of tracking the sentiments of COVID-19, negative, positive and neutral. Nevertheless, only the second one (of Nick Krystallis) can reveal specific sentiments and types of mood, such as Surprise, Disgust, Anger, Happiness, Fear and Sadness due to its special structure.

Based on what we have already discussed in the Introduction, the research questions can be shown as follows:

RQ1: The sentimental impact of COVID-19 follows the development of the pandemic.

RQ2: Vader and modified Nkryst lexicons provide similar results.

RQ3: Extracting sentimental results from the two tested electronic systems are comparable in computing resources.

The rest of this paper is structured as follows. In Section 2 we present the method and tools (data, lexicons and method of analysis), while in Section 3 the results are shown in regard to the negative of positive sentiments. Section 4 contains the discussion on the findings described before and in Section 5 we show the conclusions and the outlook of the current study.

2. Methods and tools

2.1. Data

Twitter is a social media platform with more than 500 million daily tweets worldwide and around 200 billion tweets per year (Internet live stats, 2023). We obtained data from this social network including 153,528 tweets from 27.09.2020 until 26.09.2021. The selection of this period was driven by the period of application of restricted or confinement measures among the population, such as closing of borders or businesses, working from home, restriction of movements, mask mandate or vaccination suggestion or even obligation for some specific age groups (e.g., over 60 years of age).

The tweets were obtained by using a specific library of the Twitter Streaming Application Programming Interface (API) named Tweepy (Tweepy.org, 2021; Hasan et al., 2018) written in the Python programming language (Tsakalidis et al., 2018). On the other hand, the official epidemiological data for COVID-19 was found from the European Center for Disease Prevention and Control-ECDC (ECDC, 2021). Both data sets were obtained real-time in an integrated software running as a computer application in a computer server 24/7.

ECDC, as well as other health agencies or organizations, such as the World Health Organization (WHO), the American Centers of Disease Control and Prevention (CDC), etc. publish data mainly on weekly basis (e.g., Influenza, World Health Organization-WHO, 2021) or monthly basis (e.g., Measles, ECDC, 2023), as this is the official epidemiological method of analyzing the data. However, in the case of COVID-19, through ECDC either weekly or daily data can be found; the weekly data include all countries of the world, while the daily only the countries of the European Union. We gathered weekly data because weekly monitoring is the most common method, and we wanted to have epidemiological data for all countries.

Regarding the data structure, Twitter messages include the day and time, the user, and the text, while the data from ECDC have COVID-19 weekly cases per 100,000 people.

Lexicon.

To be able to track all the words in all the previously mentioned grammatical types we have built modified lexicons and tested in two computational systems: One built in Python by using the Vader tools and another one written in VB.NET for NetCore 3.1 (Microsoft, 2022), which is cross-platform and, along with Net Framework 6.0 (Microsoft, 2021), is at least twice as fast as the previous Net frameworks for Windows of 4.7 or previous (Microsoft, 2020).

Two lexicons were used, one for the Python script with the Vader library and another one that it was used by Nick Krystallis (Nkryst). Both lexicons include the term (the key word) and its sentimental ranking, called Polarity. The Vader system has the Compound Polarity, which is the overall sentimental ranking and the individual elements of Positive, Negative and Neutral ranking. NKryst lexicon consists of the following:

The term.

The word type in grammar (Verb, Noun, Adjective, etc.).

The overall fixed polarity ranking (positive, negative and neutral).

The sentimental rankings (mood) of six types: Anger, Disgust, Fear, Happiness, Sadness and Surprise.

The specific indication of the grammar types is shown in the following table (Table 1):

We used the same polarities and sentiment ranking of the original lexicon, but we performed an extension of it, in order to include more words. To extend this lexicon, we based on the term theme or morpheme (greek-language.gr, 2008). The morpheme in morphology and linguistic theory in general refers to the minimal component of the word that has meaning or grammatical function. The study of formats begins systematically with the American Construction and its main representative (Bloomfield, 1933; Suping, 2017). From this minimal component, all other words are derived with the other word variations, such as persons (first, second and third), grammatical numbers (singular and plural), tenses (present, past and future), voices (active and passive), cases (nominative, accusative and genitive) and grammatical moods (indicative, interrogative, imperative, subjunctive, injunctive, optative, and potential).

Table 1
Grammar types.

#	Types	Explanation
1	VERB	Verb
2	NOUN	Noun
3	ADJ	Adjective
4	PRON	Pronoun
5	ADV	Adverbe
6	CONJ	Conjuction
7	INJ	Injective, exclamation
8	PART	Participle
9	PTCL	Particle
10	OTHER	Synthetic/prefixes

To be able to track the morpheme, we must isolate it from its suffix and, sometimes, from the prefix. Suffixes, in general, rely on the all the above-mentioned grammar types of the word, except from some types such as the particles, conjunctions, adverbs, etc. The mechanics and rules of isolating each word from its suffixes can be described as in the following table (Table 2):

By creating the modified lexicon, although the words listed are of the same number (2,315), we managed these words to totally capture 109,121 words of the official Greek dictionary for the Greek lexicon and 94,240 words for the translated Vader lexicon, that means almost 40 times more than the original one. This procedure can create more detailed and precise sentimental results, by simulating a large part of the official dictionary of about 20% (of the total number of the Greek words, instead of 0.40% of the original lexicon. Furthermore, this lexicon can be used in every programming IDE (integrated Development Environment) in a simply way to detect words inside any text, without losing its potential and the key elements initially provided by the developers of it.

2.2. Analysis

Our analysis consists of comparisons of the two lexicons and electronic systems. First, we performed the polarity comparison between Vader and Nkryst lexicons. This includes the following:

- **Total Polarity**: It is the overall score as calculated by the two different electronic systems. For Vader it is called *compound polarity* and for Nkryst *Fixed polarity*. These values were transformed in percentage (from 0 to 100%) to eliminate the scale differences and, therefore, to be able to be compared to each other. The transformation formula is given by the following equation

$$p = P_n / \max(P_1 \dots P_{53})\% \tag{1}$$

where p = the calculated polarity, Pn = the polarity of each week and max (P1… P53) % is the maximum value observed in the entire time series. Similar technique is used by Google for the results on the search volume on terms submitted in their search engine (Google Trends, 2021).

Polarity vs COVID-19 cases: This is performed to depict the previous values against the development of COVID-19 in Greece, as this is measured by the official weekly cases.

Positive and negative polarity: Because both systems and lexicons can express two different values for positive and negative polarity, it is important to show their values and their average deviation to show how much bigger (or lower) are the positive sentiments related to the negative ones. We may at first find awkward that a pandemic would show positive sentiments, but this may have its explanation as it will be described in the results Section.

The average deviation of the positive values as related with the negative is calculated by using the following equation

$$dev = \frac{1}{52}\sum_{i=1}^{52}\frac{|n_i|}{P} - 1 \tag{2}$$

where dev = is the average deviation of the negative values, n the negative value of each week and p is the negative value respectively.

- **Sentiments:** There is a separate calculation of the six sentiments measures by the second system: *Surprise*, *Disgust*, *Anger*, *Happiness*, *Fear* and *Sadness*

Table 2
Rules to determine the suffixes.

#	word type	has suffix	gender/voice	suffices	char length	grammar type
1	ADJ/NOUN	yes	male	“-ος,-ός,-ης,-ής,-ας,-άς”	last 2 characters	first person, singular number, nominative case
			female	“-α,-ά,-η,-ή”	last one character	first person, singular number, nominative case
			neutral	“-ι,-ί,-ο,-ό”		first person, singular number, nominative case
			irregular types	“-ώς,-ού,-ούς”, e.t.c	last one or 2 characters	first person, singular number, nominative case
2	PART	yes	male	“-ων,-ών,-ος”	last 2 characters	singular number, present tense
			female	“-ουσα,-ούσα,ή”	last one or 4 characters	singular number, present tense
			neutral	“-ούν,-ο”	last one or 3 characters	singular number, present tense
3	VERB	yes	active	“-ω,-ώ”	last one character	singular number, present tense
			passive	“-μια,-ομαι”	last one or 3 characters	singular number, present tense
4	PRON	yes	various	various	usually the last character	various grammatical numbers
5	ADV	no				
6	INJ	no				
7	PTCL					
8	CONJ					
9	OTHER	no (synthetics or prefixes)				

Notes:
Although some Greek words have some other suffixes, such as “-ις, -υς” for words of a male gender, we do not include these, as these are not found in this lexicon and are mainly archaic words.
Infinitives are not included as they are not present in the original lexicon of Krystallis.

- **Statistical correlations:** to test how the values of Vader and Nkryst systems are related to the real COVID-19 cases, we performed two statistical correlation calculations: one for the polarity data to the real cases and second of the polarity data to the volume of tweets. The following equation shows in detail these calculations:

$$R - \frac{\sum_{t=1}^{n}(X_t - \overline{X})(Y_t - \overline{Y})}{\sqrt{\sum_{t=1}^{n}(X_t - \overline{X})^2 \sum_{t=1}^{n}(Y_t - \overline{Y})^2}} \tag{3}$$

where Rxy is the Pearson Correlation Coefficient, Y are the real COVID-19 cases for the first measurement or the Twitter volume for the second measurement, y stands for the mean (expected) values of COVID-19 cases, while X represents the values of polarity for cases and × is the mean (expected) polarity values. These measurements were made to investigate the relation of the statistical distributions of the pandemic with the sentiment development within the same time frame of one year, meaning that, if a correlation is high, the bigger increase can be observed in the development of a pandemic and the more sentiments can be found (negative at most and perhaps more positive) and vice versa and with the same rate.

Technical measurement: Finally, we measured the execution time of the applications written in Python and VB.NET to compare which is the faster. This was done by applying the methods available in each programing language, for Python version 3.9.5 and NetCore.

At this point, as far polarity and the Vader system are concerned, we should distinguish two terms: the fixed polarity and the compound polarity. The first depicts the metrics of a specific type of sentiment, e.g., positive, negative or neutral, while the latter concerns the overall score, which is measured by this algorithm. In most cases it is not the same, as the second one is measured regardless of the first, while a specific sentence may have multiple measurements, e.g., a neutral polarity greater than 0 and a compound polarity greater than 0, but with different rates. The following output from the Vader system exactly reveals this for the sentence “Today is a nice day” or in Greek “Σήμερα είναι μια ωραία ημέρα”:

3. Enter sentence:

Σήμερα είναι μια ωραία ημέρα.
Sentence Overall Rated As Positive.
{’neg’: 0.0, ’neu’: 0.563, ’pos’: 0.437, ’compound’: 0.4767}.
Enter sentence:

According to Hutto and Gilbert (2014), the Compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized. Its value is rated from −1 to 1 (extreme negative to extreme positive) and can be calculated as of the following equitation:

$$x = \frac{x}{\sqrt{x2 - a}} \tag{4}$$

where:

x = is the sum of valence scores of the constituent words and.
α = the normalization constant (default value = 15).

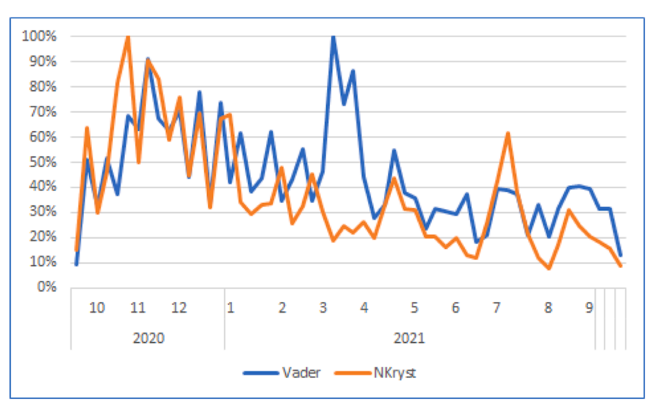

Fig. 1. Compound vs. Fixed Polarity.

4. Results

4.1. Total Polarity

As shown in Fig. 1, the polarity distributions are different after March until April of 2021. The same can be noticed in the next comparisons. Indeed, there is a very big difference between the two different lexicons. Vader shows high values during the 11th week of the year 2021, which is the year's highest (100%), while the other lexicon shows only 24.74%.

This figure shows the comparison of polarity of both lexicons used. Indeed, both lexicons show a very similar development until January 2021, where the results follow a similar distribution, while from February until mid of April of the same year, there are significance differences. The second lexicon (Nkryst) shows a very high polarity during March of around four times more than the first one (vader). This is important, given that the first lexicon can show positive or negative sentiments, while the second one can extend this analysis and reveal the real attitude of people (such as anger, disgust etc.), as it will be shown below.

4.2. Polarity and COVID-19

The analysis of polarities versus COVID-19 cases showed that the first wave of COVID-19 (November 2020) can be captured with both systems, but the second one (March 2021) can be partly captured only with Vader two weeks before (in March at the 11th week of the year, instead of April), while the third one (August 2021) cannot be identified by any of these two systems. These observations can be seen in the Fig. 2.

The above observation may be weird, but if we look at the interest of COVID-19 over time, we can see that, maybe for the first time in epidemics, this interest keeps decreasing after a certain time point, i.e., November of 2020. This strange behavior of people can be either be shown by the Twitter messages volume we have gathered from Greece and by the Google Trends for the whole world by using the search term "COVID-19", as showed in Fig. 3.

The correlation coefficients found for these two polarities sets of values, can be shown in the following table (Table 3):

From the above table, the correlations in both systems are high and significant at significance level of $p < 0.01$. This means that, the sentiments are well calculated according to the Twitter data. What is more interesting is that, as much the cases increase, the more the negative sentiments are observed. This is denoted by the negative values of the coefficients.

4.3. Positive and negative polarity

As may be expected, in an adverse and unprecedented event, such as the pandemic of COVID-19, the values of the negative polarity are much higher than those of the positive, as is shown in Fig. 4. Furthermore, we notice the following:

Both systems show some of sort of positive sentiments, although significantly lower than the negative ones.

The maximum polarity values are different since the two metrics use different ranking.

The average deviation in Vader is around 2.56 while in Nkryst is even higher 7.75.

Fig. 4 indicates the comparison between positive and negative polarity of the lexicons. We can see that both systems show lower positive sentiments and we can observe the differences related to the ratio of the positive sentiments and the negative sentiments. The ratio of the positive sentiments (blue line) /negative (red line) is 0,330074636 in absolute value for the first lexicon (vader) and 0,113952214 for the second (Nkryst). This means that it is around one third for the first one, while is around one ninth in the second one, which is even more significant. We must notice though that, the secondary axis in the second graph on the right is scaled from 0 to −3000, while the first one on the left is scaled from 0 to −1000.

4.4. Sentiments

The sentiments extracted by using the NKryst lexicon can be shown in Fig. 5. The main findings regarding the different systems we used, can be summarized as follows:

Both systems can show negative and positive sentiments, but only the one based on the Nkryst lexicon have the capability of breaking down of specific sentiments.

The development of all sentiments follows the distribution of the fixed polarity. Disgust, Anger, Fear, Sadness can be classified as negative sentiments, and happiness is a positive one. Surprise can be either positive or negative.

Both systems show extremely negative sentiments during November of 2020, with the maximum values observed in the 44th week of 2020 for Vader and the 46th week of 2020 for the second system. This is normal as, during the weeks from 44 to 46 of the year 2020, the most COVID-19 cases were identified in Greece; 10,208, 15,918, 17,829 respectively for the 44, 45 and 46th week (in addition of 17,938 cases in the 47th week). The prevailed sentiment is "surprise" (25.32%), the less noticed sentiment is "sadness" (8.5%), while "fear" is also too low (13.03%).

The overall sentiments can be summarized in Fig. 5:

The overall portion of sentiments can be found in the following Table 4:

As shown in the above table, regarding positive and negative, the prominent type could not be other than the negative, bearing in mind that, after all, a pandemic is not a convenient situation for anyone in any case.

The execution time of these different systems are away far different. Python script took 1,438.46 s (around 24 min) to extract the results, while the system written in VB.NET for Net Core 3.1 was almost eight times faster, meaning only 217.32 s (around three and a half minutes), while, by using parallel programming and exploiting all the eight cores

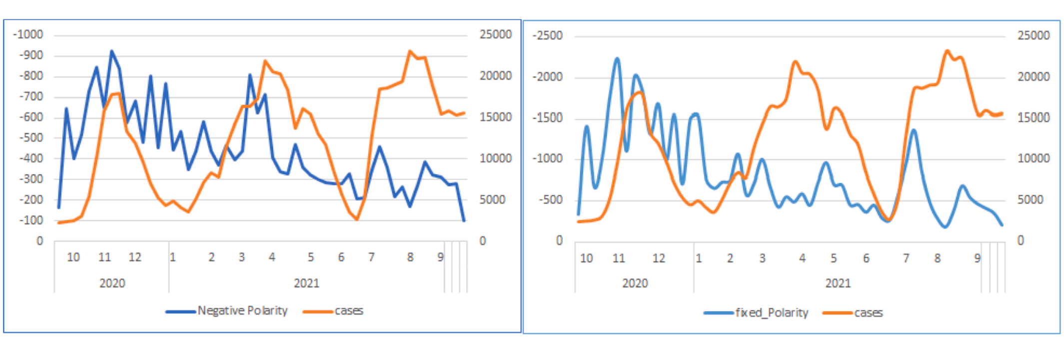

Fig. 2. Polarity vs COVID-19 cases.

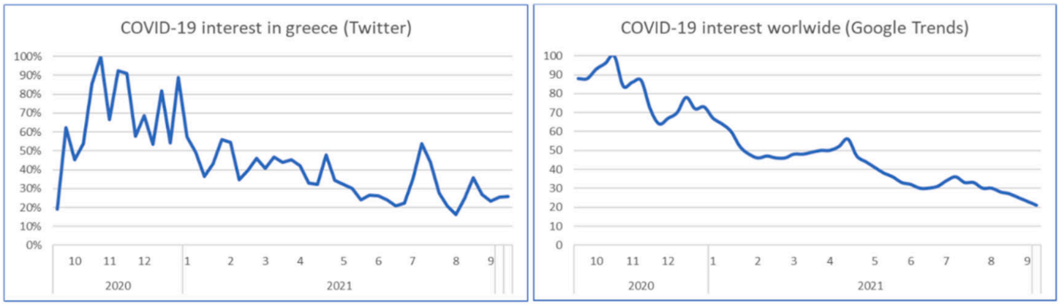

Fig. 3. COVID-19 interest over time (Greece and Worldwide).

Table 3
Correlation coefficients.

Lexicons	R	signifigant
Vader and cases	−0,70668	yes
Nkryst and cases	−0,93095	yes

and 16 threads of the Central Processing Unit (CPU), it took only 20.69 s. This means, that if we can use a more powerful machine consisting of two processors of 64 cores and 128 threads each, we could possibly perform an analysis of ten million Twitter messages in less than two minutes or around 84 s.

5. Discussion

5.1. Evidence of decreasing interest

As mentioned above, the sentimental values (positive, negative or other specific) are subjected to the development of the tweets over time. Since Internet data can reveal the interest among people, it is logical that, if this drops, all sentiments should drop as well. Particularly for the Greek case, the continuously decreasing interest over time resulted in capturing less sentiments after November of 2020, especially the negative ones. Furthermore, the interest, as found from Twitter from September of 2021, has now decreased into the level of around 20% of the total weekly interest of the period of November 2020. It' s the first time we experience a behavior like this, but this can be justified by the following reasons:

From 27 December of 2020 (National Public Health Organization, 2021), the vaccination program was initiated in Greece and until the 26th of September of 2021, around 60% (59.91%) of people have received full vaccination of two doses (emvolio.gov.gr, 2021).

From 11 of October of 2020, it was announced that no more lockdowns would be applied in the country (Kathmerini, 2021).

People may have got tired from the repeated measures and lockdowns and express less interest.

People are now experienced and have developed a strong self-awareness on how to take care and protect themselves from the virus.

Regarding the period of March of 2021, in which Vader polarity seems so high for the negative sentiments, we must note that, during this month, a new record was broken in Greece; about 70,000 cases and 1,589 deaths were reported, a lot more than the 66,020 reported cases of November 2020.

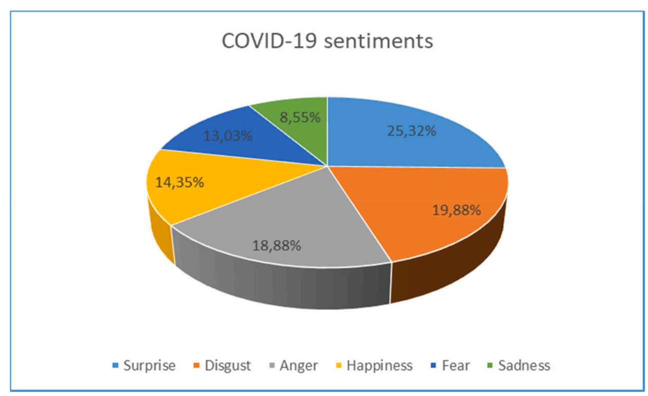

Fig. 5. COVID-19 sentiments.

Table 4
The decomposition of sentiments of each type.

Type	Vader	%
postive	21,924	13.11%
negative	49,839	29.79%
neutral	95,530	57.10%
Total	167,293	100.00%

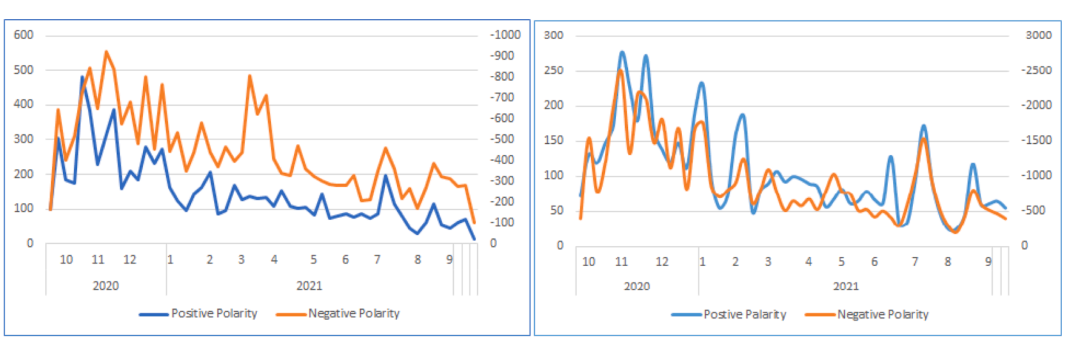

Fig. 4. Positive and negative polarity of the lexicons.

5.2. Sentiment tracking

Tracking the sentiments through the Web is possible providing that there is an appropriate ranking system and a mechanism to evaluate events reported on the Web (blogs, social media, etc.). Our research contribution lays on the following:

For the first time for Greece there are specific results for a period of an entire year.

The proposal for extending a lexicon to track about 20% of the total words of the Greek vocabulary may lead to get more reliable data samples for sentimental analysis.

The comparison of two different systems, regarding their technical characteristics and sentiment ranking.

In addition, unlike other cases of COVID-19 in other countries (e.g., Singapore), the positive polarity found in Greece is even lower than other countries. From the Vader system, they are less than one third of the total tweets in Singapore according to the work of Ridhwan and Hargreaves (2021), while from the other lexicon the positive is below the 1/8 of the total.

We agree with Garcia and Berton (Garcia and Berton, 2021) that is important to understand public reactions, information dissemination and consensus expressed by all forms and means (such as the social media) across all countries, but this may have a significant impact or influenced by critical Government interventions or decisions related to public health and also a phycological impact to individuals because of these decisions and interventions, especially for negative sentiments, such as anger or fear (Brooks et al., 2020).

Concerning the results, comparing to the findings on the recent work of Imran et al. (2022), despite we use much less data since Greece is a much smaller country, we conclude to the same result for the association of positive and negative sentiments; that the negative ones are at least twice more. In the above-mentioned study, the researchers examined almost two billion tweets and found that 52.31% are negative, while the positive are only 13.92% and ours are 29.79% and 13.11% respectively. Although our rate is smaller, the negative ones are almost the same (around 13%).

Our study includes measurements after the first lockdown in Greece in March 4th 2020. Another work by Barkur and Vibha (2020) of provides evidence from the national lockdown in India from March 25th to March 28th 2020. A total of 24,000 tweets were considered for the analysis. While the amount of the Twitter messages is a lot less, these are focused only on four days, meaning the first days of the first lockdown. The results of this study revealed that the positive sentiments were the most, as such people were expecting such a measure as lockdown to contain the spread of the virus. The positive sentiments were more than twice of the negative ones (total score around 23,000 versus less than around 10,000 of the negative), while “trust” was shown to prevail (around 17,000). “Disgust” was the lower sentiment score with around 3,000, while “surprsise” was next with about 3,500 and “sadness” follow with 6,500. In any case, these results are the opposite of those found in Greece.

5.3. Limitations: lexicons and privacy

By taking advantage of the morphology and linguistic theory, we can construct rules and computer systems to capture the sentiments inside a text. The next step and, as it seems the most difficult part of these procedures, is to aggregate the semantic orientation, or estimate values of individual words found in a text. And, according to work of Taboada (Taboada, 2016), there is an important way to do so: the Discourse patterns, but in relation to an extending procedure through finding relations of concession and conditions as proposed by some theories, e.g., a condition relation will limit the extent of a positive evaluation (Mann and Thompson, 1988). No matter what language we use, there may be similar rules of evaluating a text or individual words, but as we show for the Greek language, it may require additional patterns for each different language. In any case, we think that if we can improve lexicons to track as many as possible words of a certain language vocabulary, the most reliable data samples we may get out of this procedure, since we can take advantage of a bigger part of this language vocabulary. However, the proper selection of the key words of phrases towards a sentimental tracking remains critical.

Apart from the lexicon building, another issue refers to public health and privacy. What social media have as advantage is at the same time a drawback: the ability to identify, not only the general trend, but also specific personal information, such as the username, the message itself and the location which this comes from, at least whenever this information is available. The capability of collecting mass personal information from electronic systems for individual identification, advertisement, and tracking through multiple channels, like wearable devices, microphones, heart and respiratory monitors, and user interactions is something that nowadays is observed to the extent that we have never seen before (Data Privacy Manager, 2021). While public health can certainly benefit from electronic systems of this kind, according to the American Psychological Association, the danger of extracting personal data could result in providing an intimate window into a person’s life by revealing, not only one person’s particular movements, but through them his familial, political, professional, religious, and sexual associations by constructing individual profiles possibly for non-health purposes (Robbennolt, 2020).

Aside from the above, as regards the knowledge-based ruling and methods, we could lay down two remarks, as already discussed in the work of by Pantic et al.:

The first has to do with the linguistic analysis and rules. For example, as we presented in the analysis Section (Section 2.3), a knowledge base can correctly classify the sentence “This is a nice day” as being happy and positive, but it is attainable or possible to fail on a sentence like “Today wasn’t a nice day at all”.

In some cases, the opposite meaning of an unhappy day, as in the previous paragraph, could not lead into a negative sentiment (as it might be expected) and this is the reason of so many neutral sentiments, observed by us and by other researchers as well. On the other side, while Pantic et al. propose the concurrent use of language-based techniques along with statistical processing, even then it is feasible to achieve a dissolution of the sentiments, enough to reveal the impact of a certain situation to the people, such as the case of a pandemic.

For instance, if we simply put into Vader the sentence “Today is not a nice day” in Greek (“σήμερα δεν είναι ωραία μέρα”, we get a neutral polarity of 0.563 and a positive of 0.437. Also, if we enter the sentence “today is not a nice day at all” in Greek (“Σήμερα δεν είναι καθόλου μια ωραία μέρα”) the result is not also the expected one with a polarity of neu’: 0.659, ’pos’: 0.341. It seems that the use of the negative verbs or particles (“is not”, “do not”) may cause this effect.

Conclusions and outlook.

Sentiment tracking has been of a great interest during the last years of research. With our work we tried to deal with specific problems, regarding the Greek language and the development of Twitter messages in a long period of one year within the pandemic.

By using, extending and comparing two different implementations and lexicons to extract sentiments of the ongoing COVID-19 pandemic, we found that, as expected, negative sentiments in Vader dominate almost three times more than the positive ones and about nine times more when using the Nkryst modified system. Moreover, the polarity calculated by both systems does not follow the exact development of the COVID-19 pandemic, since the interest of people for COVID-19 has dropped to almost 20% of what it was observed during the year of 2020, despite that COVID-19 had three major waves from late September of 2020. This happens because, if we correlate the data volume (Twitter messages or possibly Google search engine data) to the epidemiological data (COVID-19 cases), the correlation coefficient is very low (R2 is 0,167387), although it seems higher during the first four months of the examination period from September 2020 until January 2021, i.e., before the interest of the time drops a lot. This is the opposite of what was found in similar other cases when correlating just the volume of the data from the Internet, e.g., Twitter with epidemiological data on various infectious diseases, e.g., influenza, malaria, scarlet fever or measles. Fortunately, this does not cause the absence of the sentiment tracking.

By using sentiment analysis, we managed to evaluate the sentiments for the entire time series of a period of exactly one year and found that negative emotions were the most frequent, as an epidemic situation is doubtful to cause the opposite. Concretely, surprise and disgust were the prevalent sentiments.

Although we did not change the total number of the words inside the lexicons we used, we managed these words to have the potential to capture totally at least around 100,000 words of the official Greek dictionary with different grammatical types or patterns.

As mentioned in the Discussion Section, the most difficult task is to be able to evaluate each word or sentence or even an entire text inside a text and this should be done as fast as possible, given that the data from the Twitter could concern bigger countries than Greece with al lot of more messages. The Python script took totally much more time (almost 24 min, although in Linux the total time was reduced by around 40%) to extract results than the system written for NET Core environment (almost three and a half minutes for a single thread application and around twenty seconds by a multithread algorithm).

Finally, it should be considered that sentiments are generally influenced by the individual situation of a person, but also by the external environment, such as the health conditions or the policy implementation by Governments or health organizations, while some policies could result in a severe phycological impact among people. We should not forget the proposition of the recent research by Huang et al. (2022), that mining and analyzing social networks should be done in a quick and immediate way in order to accomplish early warning and detection, which in turn could lead in human mobility monitoring by measuring public attitudes and emotions and, consequently, avoiding misinformation, detestation, social unrest and possible violent behaviors. This and many previous studies discuss the sentiment analysis of one given epidemic or pandemic, such as COVID-19. Towards a possible future work, it would be interested to extend this analysis about the occurrence of different types of infectious events, such as epidemics, pandemics, viruses or outbreaks, during a major time frame, e.g., during the last 10 years, in order to retrieve resemblances or differences (Alamoodi et al., 2021). Furthermore, it would also be interesting to perform a comparative study on the main models used so far; the lexicon-based, machine learning-based, hybrid-based and individuals' models.

In another direction, as previous said in the Discussion section, a hard task is the aggregation of the semantic orientation, because values of individual words found in a text are subjected to errors or difficulties to be detected. Such cases are the detection of irony and sarcasm. Therefore, it would be very interesting to use some methods to detect these by using Feature Fusion and Ensemble Classifier (Keerthi Kumar & Harish, 2019) or with other specific language resources, such as morphological dictionaries, sentiment lexicon, lexicon of markers and a WordNet based ontology (Mladenovic, 2017). We could also explore the concepts of unexpectedness and contradiction that seems to be frequent in expressions (Potamias et al., 2020).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agarwal A., Xie B., Vovsha I., Rambow O., Passonneau R. (2011), Sentiment Analysis of Twitter Data, Proceedings of the Workshop on Language in Social Media (LSM 2011), 30–38, Portland, Oregon, 23 June 2011, c 2011 Association for Computational Linguistics, https://aclanthology.org/W11-0705.pdf.

Alamoodi, A. H., Zaidan, B. B., Zaidan, A. A., Albahri, O. S., Mohammed, K. I., Malik, R. Q., Almahdi, E. M., Chyad, M. A., Tareq, Z., Albahri, A. S., Hameed, H., & Alaa, M. (2021). (2021), Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Systems with Applications, 167*, Article 114155. https://doi.org/10.1016/j.eswa.2020.114155, PMID: 33139966; PMCID: PMC7591875

American Centers of Disease Control and Prevention (CDC), https://www.cdc.gov/ (last accessed 29.04.2023).

Arbane M., Benlamri R., Brik Y., Alahmar A.D. (2023), Social media-based COVID-19 sentiment classification model using Bi-LSTM, Expert Systems with Applications 212 2023, 118710 ISSN 0957-4174, doi: 10.1016/j.eswa.2022.118710.

Barkur, G., & Vibha, K. G. B. (2020). Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India. *Asian Journal of Psychiatry, 51*, 02089. https://doi.org/10.1016/j.ajp.2020.102089. Epub 2020 Apr 12. PMID: 32305035; PMCID: PMC7152888.

Bartneck C., Lyons M.J., Saerbeck M. (2008), The Relationship Between Emotion Models and Artificial Intelligence, In Proceedings of the SAB 2008 Workshop on The Role of Emotion in Adaptive Behavior and Cognitive Robotics, Osaka, arXiv:1706.09554 [cs. HC], doi: 10.48550/arXiv.1706.09554.

Boon-Itt, S., & Skunkan, Y. (2020). Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study. *JMIR Public Health and Surveillance, 6*(4), e21978. PMID: 33108310, PMCID: 7661106.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet, 395*(10227), 912–920. https://doi.org/10.1016/S0140-6736(20)30460-8

Catapang J.K. and Cleofas J.V. (2022), Topic Modeling, Clade-assisted Sentiment Analysis, and Vaccine Brand Reputation Analysis of COVID-19 Vaccine-related Facebook Comments in the Philippines, in IEEE 16th International Conference on Semantic Computing (ICSC) 2022, 123-130, doi: 10.1109/ICSC52841.2022.00026.

Chen N., Chen X., Pang J. (2022), A multilingual dataset of COVID-19 vaccination attitudes on Twitter, Data in Brief 44 Science Direct 2022 (4), ISSN 2352-3409, doi: 10.1016/j.dib.2022.108503.

Chen, C., Zhang, Q., Kashani, M. H., Jun, C., Bateni, S. M., Band, S. S., Dash, S. S., & Chau, K.-W. (2022b). Forecast of rainfall distribution based on fixed sliding window long short-term memory. *Engineering Applications of Computational Fluid Mechanics, 16*(1), 248–261.

Data Privacy Manager (2021), What is the future of your Privacy in METAverse, article 12.12.2021, https://dataprivacymanager.net/what-is-the-future-of-your-privacy-in-facebook-metaverse (last accessed 12.02.2023).

emvolio.gov.gr (2021), Covid-19 Vaccination statistics, https://emvolio.gov.gr/vaccinationtracker (last accessed 12.02.2023).

European Centre for Disease Prevention and Control-ECDC (2023), Monthly measles and rubella monitoring, https://www.ecdc.europa.eu/en/rubella/surveillance-and-disease-data/monthly-measles-rubella-monitoring-reports (last accessed 12.02.2023).

European Centre for Disease Prevention and Control-ECDC (2021), COVID-19, https://www.ecdc.europa.eu/en/covid-19 (last accessed 12.02.2023).

Garcia K., Berton L. (2020), Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA, Applied Soft Computing 101 107057, ISSN 1568-4946, doi: 10.1016/j.asoc.2020.107057.

Giachanou A, Crestani F. (2017), Like It or Not: A Survey of Twitter Sentiment Analysis Methods, ACM Computing Surveys 49 (2) June 2017 Article No. 28, 1–41, doi: 10.1145/2938640.

Github (2020), greek-sentiment-lexicon, https://github.com/MKLab-ITI/greek-sentiment-lexicon (last accessed 12.02.2023).

Github (2021), titoBouzout/Dictionaries, Hunspell UTF8 dictionaries. https://github.com/titoBouzout/Dictionaries (last accessed 12.02.2023).

Github (2022), Greek-Sentiment-Analysis, https://github.com/NKryst/Greek-Sentiment-Analysis#readme (last accessed 12.02.2023).

Gitub, cjhutto vaderSentiment (2022), Valence Aware Dictionary and sEntiment Reasoner, https://github.com/cjhutto/vaderSentiment (last accessed 12.02.2023).

Google Trends (2021), Trends, FAQ about Google Trends data, https://support.google.com/trends/answer/4365533?hl=en (last accessed 12.02.2023).

greek-language.gr (2008), The gate for the Greek language, μόρφημα [morpheme], https://www.greek-language.gr/greekLang/modern_greek/tools/lexica/glossology/show.html?id=61 (last accessed 12.02.2023).

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications., 23*(1), 11. https://doi.org/10.3390/mca23010011

He, L., & Zheng, K. (2019). How do general-purpose sentiment analyzers perform when applied to health-related online social media data? *Studies in health technology and informatics, 264*, 1208–1212. https://doi.org/10.3233/SHTI190418

Hongjie, D., Daji, E., Fangyao, L., Ying, C., & Bo, M. (2022). Text sentiment analysis of fusion model based on attention mechanism,. *Procedia Computer Science, 199*. https://doi.org/10.1016/j.procs.2022.01.092. ISSN 741–748.

Huang, X., et al. (2022). Social media mining under the COVID-19 context: Progress, challenges, and opportunities. *International Journal of Applied Earth Observation and Geoinformation, 113*, 102967. https://doi.org/10.1016/j.jag.2022.102967. ISSN 1569–8432.

Hutto, C., & Gilbert, E. (2014), VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text, In Proceedings of the International AAAI Conference on Web and Social Media 8(1), 216-225, doi: 10.1609/icwsm.v8i1.14550.

Imran, M., Qazi, U., & Ofli, F. (2022). TBCOV: two billion multilingual COVID-19 Tweets with sentiment, entity geo, and gender labels. *Data, 7*(1), 8. https://doi.org/10.3390/data7010008

Internet live stats (2023), Trends and More, Twitter Usage Statistics, https://www.internetlivestats.com/twitter-statistics (last accessed 12.02.2023).

Kalamatianos, G., Mallis, D., Symeonidis, S., & Arampatzis, A. (2015). Sentiment analysis of greek tweets and hashtags using a sentiment lexicon, In PCI '15. In *Proceedings of the 19th Panhellenic Conference on Informatics October*. https://doi.org/10.1145/2801948.280201

Kathmerini (2021), Has the risk of a new lockdown in our country passed?, https://www.kathimerini.gr/economy/561534250/echei-parelthei-o-kindynos-neoy-lockdown-sti-chora-mas/ (last accessed 12.02.2023).

Keerthi Kumar, H. S., & Harish, B. S. (2019). Automatic irony detection using feature fusion and ensemble classifier, international journal of interactive multimedia and artificial. *Intelligence, 70–79*. https://doi.org/10.9781/ijimai.2019.07.002

Kruspe A., Häberle M., Kuhn I., Zhu X.X. (2022), Cross-language sentiment analysis of European Twitter messages during the COVID-19 pandemic, Computer Science Social and Information Networks, arXiv:2008.12172 [cs.SI], doi: 10.48550/arXiv.2008.12172.

Kydros D., Argyropoulou M., Vrana V. (2021), A Content and Sentiment Analysis of Greek Tweets during the Pandemic, Sustainability 13(11)6150, doi: 10.3390/su13116150.

Lo, S. L., Cambria, E., Chiong, R., et al. (2017). Multilingual sentiment analysis: from formal to informal and scarce resource languages. *Artificial Intelligence Review, 48*, 499–527. https://doi.org/10.1007/s10462-016-9508-4

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Interdisciplinary Journal for the Study of Discourse, 8*(3), 243–281. https://doi.org/10.1515/text.1.1988.8.3.243

Microsoft (2020), Microsoft Net Framework, https://dotnet.microsoft.com/download/dotnet-framework (last accessed 12.02.2023).

Microsoft (2021), Announcing .NET 6 — The Fastest .NET Yet, https://devblogs.microsoft.com/dotnet/announcing-net-6 ((ast accessed 12.02.2023).

Microsoft (2022), Microsoft Net Core, https://dotnet.microsoft.com/en-us/download/dotnet/3.1, (last accessed 12.02.2023).

Minsky, M.L. (1986), The society of mind. Simon and Schuster, New York N.Y, Published by Simon & Schuster, Inc. Simon & Schuster Building Rockefeller Center IZJO Avenue of the Americas New York, NY 10020, ISBN 0-671-60740-5, rsBN 0-671-6571 l-5 Pbk.

Mladenovic M., J. Krstev J., Mitrovic J., Stankovic J. (2017), Using Lexical Resources for Irony and Sarcasm Classification, in BCI '17: Proceedings of the 8th Balkan Conference in Informatics September 2017 Article No. 13,1–8, doi: 10.1145/3136273.3136298.

National Public Health Organization (2021), Press release, https://eody.gov.gr/wp-content/uploads/2021/01/20210111.pdf (last accessed 12.02.2023).

Okango, E., & Mwambi, H. (2022). Dictionary sus (COVID-19) effects and response. *Annals of Data Science, 9*, 175–186. https://doi.org/10.1007/s40745-021-00358-5

Orduz J.C. (2018), Text Mining, Networks and Visualization: Plebiscito Tweets, https://juanitorduz.github.io/text-mining-networks-and-visualization-plebiscito-tweets ((last accessed 12.02.2023).

Pantic M.et al. (2006), Affective Multimodal Human-Computer Interaction, in Proceedings of the 13th annual ACM international conference on Multimedia November 2005, 669–676, doi: 10.1145/1101149.1101299.

Potamias, R. A., Siolas, G., & Stafylopatis, A. (2020). A transformer-based approach to irony and sarcasm detection. *Neural Computing and Applications, 32*, 17309–17320. https://doi.org/10.1007/s00521-020-05102-3)

Ridhwan, K. M., & Hargreaves, C. A. (2021). Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore. *International Journal of Information Management Data Insights, 1*(2), 100021. https://doi.org/10.1016/j.jjimei.2021.100021. ISSN 2667–0968.

Robbennolt J.D. (2020), Privacy versus public health, American Psychology Association 51 No. 4 "Judicial Notebook", https://www.apa.org/monitor/2020/06/jn (last accessed 12.02.2023).

Samaras L., García-Barriocanal E., Sicilia M.A. (2020), Syndromic surveillance using web data: a systematic review, Innovation in Health Informatics, 2020, 39–77. doi: 10.1016/B978-0-12-819043-2.00002-2. Epub 2019 Nov 15. PMCID: PMC7153324.

Suping D. (2017), Bloomfield's concept of meaning, Journal of Language Teaching and Research 8 (2), Academy Publication Co., LTD, ISSN 1798-4769, eISSN 2053-0684.

Swathi, T., Kasiviswanath, N., & Rao, A. A. (2022). An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis. *Applied Intelligence, 52*, 13675–13688. https://doi.org/10.1007/s10489-022-03175-2

Taboada M. (2016), Sentiment Analysis: An overview from Linguistics, Annual Review of Applied Linguistics, https://www.researchgate.net/publication/283954600_Sentiment_Analysis_An_Overview_from_Linguistics#pf1f (last accessed 12.02.2023).

Tsakalidis, A., Papadopoulos, S., Voskaki, R., et al. (2018). Building and evaluating resources for sentiment analysis in the Greek language. *Lang Resources & Evaluation, 52*, 1021–1044. https://doi.org/10.1007/s10579-018-9420-4

Tweepy.org (2021), Twitter API v1.1 Interface, https://docs.tweepy.org/en/stable/api.html (last accessed 12.02.2023).

World Health Organization-WHO (2021), Weekly Epidemiological Record (WER), https://www.who.int/publications/journals/weekly-epidemiological-record (last accessed 12.02.2023).

Yadav, V., Verma, P. & Katiyar, V. Long short term memory (LSTM) model for sentiment analysis in social data for e-commerce products reviews in Hindi languages (2023) Int. j. inf. tecnol. 15, 759–772 (2023), doi: 10.1007/s41870-022-01010-y.

Yajun, F., Kangkang, X., Hui, W., Ying, Z., & Bo, T. (2020). Spatiotemporal modeling for nonlinear distributed thermal processes based on KL decomposition, MLP and LSTM Network. *IEEE Access., 8*, 25111–25121.

Further reading

Banan, A., Nasirii, A., & Taheri-Garavand, A. (2020). Deep learning-based appearance features extraction for automated carp species identification. *Aquacultural Engineering, 89*(102053). https://doi.org/10.1016/j.aquaeng.2020.102053

Xiao, X., et al. (2022). Social media mining under the COVID-19 context: Progress, challenges, and opportunities. *International Journal of Applied Earth Observation and Geoinformation, 113*, Article 102967. https://doi.org/10.1016/j.jag.2022.102967