Contents lists available at ScienceDirect

Applied Soft Computing Journal

Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA

Klaifer Garcia, Lilian Berton *

Institute of Science and Technology, Federal University of Sao Paulo, São José dos Campos, São Paulo, 12247-014, Brazil

ARTICLE INFO

ABSTRACT

Twitter is a social media platform with more than 500 million users worldwide. It has become a tool for spreading the news, discussing ideas and comments on world events. Twitter is also an important source of health-related information, given the amount of news, opinions and information that is shared by both citizens and official sources. It is a challenge identifying interesting and useful content from large text-streams in different languages, few works have explored languages other than English. In this paper, we use topic identification and sentiment analysis to explore a large number of tweets in both countries with a high number of spreading and deaths by COVID-19, Brazil, and the USA. We employ 3,332,565 tweets in English and 3,155,277 tweets in Portuguese to compare and discuss the effectiveness of topic identification and sentiment analysis in both languages. We ranked ten topics and analyzed the content discussed on Twitter for four months providing an assessment of the discourse evolution over time. The topics we identified were representative of the news outlets during April and August in both countries. We contribute to the study of the Portuguese language, to the analysis of sentiment trends over a long period and their relation to announced news, and the comparison of the human behavior in two different geographical locations affected by this pandemic. It is important to understand public reactions, information dissemination and consensus building in all major forms, including social media in different countries.

1. Introduction

The COVID-19 outbreak has been declared a pandemic by the World Health Organization (WHO), because of its high spreading and severity, which can cause severe pneumonia, respiratory failure, and death [1]. It was characterized as a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).[1] Many public health approaches have been adopted like hand hygiene, social distancing, and self-isolation, especially because, until now, there is no vaccine and appropriate treatment for COVID-19.

Every day, a large number of websites and online social media produce a huge amount of data. We have many types of social networks like micro-blogging platforms, blogging, instant messaging Apps, networking, software elaboration, photo/video sharing. All these social media can be an important source of data for detecting outbreaks and helping to understand public attitudes and behaviors during a crisis [2,3]. Moreover, they permit sharing information faster than textbooks or journals, which can

be critical for knowledge translation and dissemination. They also influence political communication and public debate.

Social networks and online platforms are a powerful tool for world leaders to rapidly communicate public health information with citizens. One of the most used is Twitter. According to [4], the G7 world leaders have around 85.7 million followers. Twitter is a free micro-blogging platform with 152 million registered daily users [5] and over 500 million people visit Twitter per month without logging into an account [6]. Some papers have analyzed Twitter data related to COVID-19, such as [4] which made a content analysis on viral tweets from G7 world leaders. Authors in [7] analyzed the most retweeted English-language tweets on Twitter mentioning COVID-19 during March 2020. Tweets have been analyzed regarding different diseases and disasters, like Zika [8,9], Ebola [10], the Japanese earthquake of 2012 [11], Hurricane Irma [12]. In many cases, tweets share first-hand information quickly and even tweets from citizens can reach large audiences during crises.

Some tools can be used to obtain Twitter trends, most of them are based on hashtags; however, not all tweets related to a given topic are hashtagged. To find relevant information or topics is a difficult task since there are millions of daily tweets covering thousands of topics, there is noisy vocabulary (slangs, emoticons, grammar errors), the text is very short (140 characters), and multilingual tweets. Topic detection is a technique

---

* Corresponding author.
E-mail address: lberton@unifesp.br (L. Berton).
[1] www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it.

for discovering the main topics automatically [13]. Few works explore topic detection for languages different from English, like Portuguese [14,15].

Besides many individuals, organizations and governments are using Twitter to communicate content related to COVID-19, most works focus on English tweets and short time-periods (less than one month). Not much is known about the topics oscillation during a long time-period and what is being shared in non-English tweets. Here, we analyzed tweets in English and Portuguese, mainly from the United States of America (USA) and Brazil. Both countries have the highest number of deaths by COVID-19 in the world. Until 30 October WHO published that the USA had 8,852,730 confirmed cases with 227,178 deaths, Brazil had 5,494,376 confirmed cases with 158,969 deaths.[2] Analyzing this information can help to understand people's behavior, to support communication and health promotion messaging. Moreover, machine learning (ML) approaches and natural language processing (NLP) techniques can be very useful to describe the amount of information being micro-blogged about COVID-19.

This paper aims to apply topic modeling and sentiment analysis methods on COVID-19 Twitter data comparing English and Portuguese languages. The main questions investigated here are: (i) topics from different languages (especially English and Portuguese) are distinct from each other? (ii) how is the evolution of existing topics over the temporal dimension? (iii) how is the sentiment in the topics? (iv) what are the reasons for the topics?

The main contributions of our work are summarized as follows:

- To the best of our knowledge, this is the first work performing sentiment analysis and topic detection in Portuguese tweets about COVID-19. Moreover, few papers in the literature have employed text mining tasks in Portuguese texts. So, we contribute to the study of this language.
- We also compare sentiments and topics associated with COVID-19 in the USA. This way, we consider Twitter data from two different countries with a high number of infections and deaths by COVID-19 investigating how this pandemic escalates in different geographical locations.
- We analyzed data considering four months, so it was possible to see the oscillation of topics/sentiments during a large period, and the trends in the context of its duration, frequency, and relation to around 100 news retrieved from Google news.
- We made a broad comparison among many classifiers for sentiment analysis in Portuguese and English Twitter data. We combined recent embedding models (SBert, mUSE, FastText) for feature extraction.
- We collect, pre-process, and will make available a large COVID-19 tweets dataset for Portuguese and English which may create opportunities for further studies.

Our findings suggest that most of the topics are similar in both languages (seven from ten identified topics). The actual world globalization leads all the countries passing to similar problems regarding economic impacts, COVID-19 proliferation, and treatment discussions. The negative messages dominate along with the four monthly variations of the topics. For English tweets only topics related to treatments and sports have the number of positive messages close to negative ones. For Portuguese tweets topics related to politics, treatments and sports have the number of positive messages higher than negative. The oscillation in the number of messages for each topic was influenced most by political actions and statements in both countries. We listed

much news from different fonts of information to understand the volume oscillation and the topics.

The remainder of the paper is organized as follows: Section 2 presents some related works that also employed the topic model and sentiment analysis into COVID-19 data. Section 3 presents the algorithms we employed for topic detection and sentiment analysis. Section 4 presents the dataset and the methodology employed in this work. Section 5 presents the top 10 topics identified in English and Portuguese tweets, the topic oscillation along the four months and sentiment analysis over the topics. Section 6 presents the discussion about the results. Section 7 brings the final remarks and future works.

2. Related work

Some works employed topic models in tweets to analyze different concerns about COVID-19. We summarize them below regarding the techniques employed and the main topics identified. Most of the papers explored English tweets only.

In [16] authors identify the main topics posted by Twitter users related to the COVID-19 from public English language tweets from February 2, 2020, to March 15, 2020. They employed Latent Dirichlet Allocation (LDA) for topic modeling. They identified 12 topics, which were grouped into four main themes: the origin of the virus; its sources; its impact on people, countries, and the economy; and ways of mitigating the risk of infection. The mean sentiment was positive for 10 topics and negative for 2 topics (deaths caused by COVID-19 and increased racism).

In [17] authors applied a Biterm Topic Model (BTM) in tweets collected from March 3–20, 2020 related to COVID-19 symptoms, this model separates into topic clusters groups of tweets containing the same word-related themes about symptoms, testing, and recovery. Tweets were grouped into five main categories: first- and second hand reports of symptoms, symptom reporting concurrent with lack of testing, discussion of recovery, confirmation of negative COVID-19 diagnosis after receiving testing, and users recalling symptoms and questioning whether they might have been previously infected with COVID-19. The users were not able to get tested to confirm their concerns.

Authors in [18] performed a comparative analysis on five different social media platforms (Twitter, Instagram, YouTube, Reddit, and Gab) during the COVID-19 health emergency. They clustered the data by running the Partitioning Around Medoids (PAM) and using as proximity metric the cosine distance matrix of words in their vector representations. They identify topics for each social media, in Twitter 21 topics were found on data collected from January 27 to February 14: suspended flights and repatriation, economic impact, protection advice, prayers, God blessing request, death toll, infection rates, biological warfare, communist regime, Huoshenshan hospital, comparison with other viruses, Chinese wet markets, virus spreading, disease description and symptoms, racism, other. Moreover, they model the spread of information and found that even the information reliable or questionable has similar spreading patterns.

Regarding sentiment analysis, authors in [19] examined worldwide Twitter from January 28 to April 9, 2020. The authors recovered more than 20 millions tweets. They employed a lexical approach and the CrystalFeel algorithm. Four emotions fear, anger, sadness and joy emerged and the reasons associated with them were investigated. Authors in [20] collected 226,668 tweets between December 2019 and May 2020 and compared to 23,000 most retweeted tweets from 1st January 2019 to 23 March 2020 (which has a minimum of 1000 re-tweets). They found while the number of positive and neutral tweets is high, the number of retweeted tweets is most negative.

Authors in [19] collected 20 millions worldwide English tweets in the period of January 28 to April 9, and analyzed trends of

four emotions: fear, anger, sadness, and joy using the algorithm CrystalFeel. They only generate word clouds to identify possible topics related to emotions, which suggest that fears around shortages of COVID-19 tests and medical supplies. Anger shifted from xenophobia to stay-at-home notices. Sadness related to topics of losing friends and family members, and joy included words of gratitude and good health.

A longer period of Twitter coronavirus-related data from January 28 to July 1st were collected by [21]. They annotated each tweet with seventeen latent semantic attributes related to ten detect topics by LDA algorithm and seven attributes related to sentiments retrieved by CrystalFeel algorithm. They found that anger was the dominant emotion in tweets.

A recent survey retrieved a few studies that examined the presence of pandemics such as COVID-19 via sentiment analysis [22], probably because the previous epidemics were smaller. Social media is an important news medium nowadays, studies that help understand people's behavior could help authorities manage a situation.

3. Background

Following, Section 3.1 describes the algorithm used for topic detection in English and Portuguese data. Section 3.2 describes the algorithm used for sentiment analysis in English data. The sentiment analysis in Portuguese data were performed by machine learning classification, using several input attributes, such as sentence embeddings, which were produced with Universal Sentence Encoder, described in Section 3.3.

3.1. Topic modeling

Topic modeling is a technique used to extract and summarize trending issues from documents. Among the existing techniques, one of the most common is the Latent Dirichlet Allocation [23] (LDA), which represents each topic as a probability distribution over the words in a dictionary. However, traditional topic models experience large performance degradation over short texts due to the lack of word co-occurrence information [24]. Therefore, some techniques have been developed that are better adapted to short texts, like the collapsed Gibbs Sampling algorithm for the Dirichlet Multinomial Mixture (GSDMM), which was used in this work. This technique simplifies the inference process assuming that each document is the result of a single topic [25].

As in other techniques, GSDMM supposes a generative process to build documents, where the terms/words are drawn from a probability distribution, with each probability distribution representing a topic. This generative process is represented in the Algorithm 1. In this algorithm θ is the document-topic distribution and Θ is the word-topic distribution, which is initialized with parameters α and β respectively. K is the number of topics, D is the number of documents to be generated, w are words in document d and z_d is the relationship between a document and a topic.

Training consists of adjusting the probability distributions to maximize the probability of producing a set of documents. This procedure is performed with Collapsed Gibbs Sampling, and the probability of relationship between documents and topics is calculated by Eq. (1).

$$p(z_d = z | \vec{z}_{\neg d}, \vec{d}) \propto \frac{m_{z, \neg d} + \alpha}{D - 1 + K\alpha} \frac{\prod_{w \in d} \prod_{j=1}^{N_d^w} (n_{z, \neg d}^w + \beta + j + 1)}{\prod_{i=1}^{N_d} (n_{z, \neg d} + V\beta + i - 1)}$$

(1)

In this equation, N_d is the number of words in document d, m_z is the number of documents in topic z, n_z is the number of words in topic z, n_z^w is the number of occurrences of the word w in topic z, $\neg d$ means that document d was excluded from the attribute and V is the number of words in the vocabulary.

Algorithm 1: GSDMM generative process

Sample a topic distribution $\theta \sim Dirichlet(\alpha)$;
for $k \in [1, ..., K]$ **do**
  Draw a topic-word distribution $\Theta_k \sim Dirichlet(\beta)$;
end
for $d \in [1, ..., D]$ **do**
  Sample a topic $z_d \sim Multinomial(\theta)$;
end
for $\{w_1, ..., w_{|d|}\} \in d$ **do**
  Sample a word $w_n \sim Multinomial(\Theta_{zd})$;
end

3.2. Sentiment analysis

For English messages, CrystalFeel [26] was used, which was first presented at SemEval 2018. This method uses a combination of affective lexicons, N-grams, POS (part-of-speech), word counts, and word embedding in an SVM classifier, to identify the intensity of different emotions (joy, fear, sadness and anger) and valence (positive, negative, neutral). In CrystalFeel, to relate the messages to the dominant emotion, the Algorithm 2 is applied. In this algorithm, it is possible to notice that the majority of positive messages are related to joy emotion, while negative messages are split between anger, fear and sadness.

Algorithm 2: Emotion score for label in CrystalFell

Result: Message label
emotion-category = "no specific emotion";
if *valence-score > 0.52* **then**
  emotion-category = "joy";
else
  if *valence-score < 0.48* **then**
    emotion-category = "anger";
    if *(fear-score > anger-score) and (fear-score > sadness-score)* **then**
      emotion-category = "fear";
    else
      if *(sadness-score > anger-score) and (sadness-score > fear-score)* **then**
        emotion-category = "sadness";
   **end**
    end
  end
end

3.3. Sentence embedding

Another approach that we tested for sentiment analysis is the transfer of learning from Semantic Retrieval (SR) techniques. An example is the Multilingual Universal Sentence Encoder for Semantic Retrieval [27] (mUSE). This technique produces sentence embedding with semantic meaning, which means that it is possible to find similar texts by comparing the embedding vectors.

In mUSE, the authors present two architectures for the construction of sentence embedding, one more accurate and computationally costly using a Transformer, and another not so accurate but more computationally efficient, using convolutional neural networks. The networks were trained with pairs of questions and answers, pairs of translations and the Stanford Natural Language
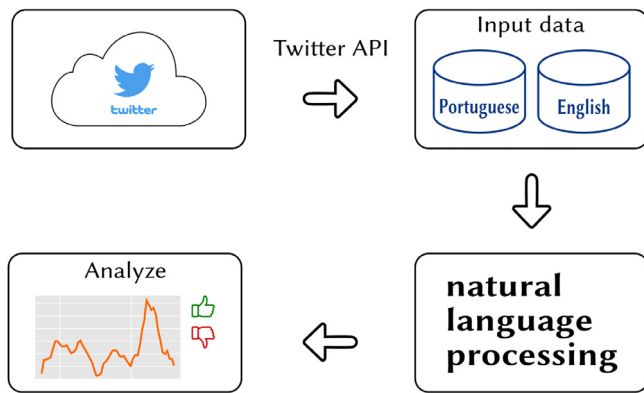
Fig. 1. Work-flow: we access Twitter API and collect tweets using two different queries, one for English and other for Portuguese. The messages are processed separately, in the natural language processing step, to be compared in the analysis step.

Inference[3] (SNLI) corpus, which is a set of pairs of sentences annotated with entailment, contradiction, and neutral. The trained models were evaluated in different Retrieval Tasks.

Another model we employed was Sentence BERT [28] (SBERT). In the SBERT, a pooling layer is added to the output of BERT [29], which is responsible for producing embedding of the sentences. This pooling layer simply performs a medium or max-over-time operation on the BERT output (CLS-token). An important contribution of this work is the BERT fine-tuning procedure, which occurs using a structure of Siamese networks. For this, two BERT networks, each with a pooling layer, receive different sentences and produce different embedding vectors. These outputs are combined in a problem dependent layer, which determines the intensity of the adjustment. An example, presented by the authors, is the comparison of the output vectors using the cosine distance, which could be optimized with mean-squared-error loss.

4. Material and methods

The methodology employed in this work is depicted in Fig. 1. The tweets were collected using two different queries, one for English and other for Portuguese. To delimit the results in a single language, we used the API filtering and a set of terms in the corresponding language. More details on the data collected can be found in Section 4.1.

In the natural language processing step, messages in both languages are processed separately. Although operations with similar objectives were applied, the codes were not exactly the same, since some of the resources used for English were not available for Portuguese. Natural language processing operations are described in more detail in Fig. 2. Then, after running topics and sentiment analysis algorithms the results are analyzed.

As shown in Fig. 2, all operations were applied to the message texts. The first operation is preprocessing which includes operations that are common for topic modeling and sentiment analysis. However, some methods may require additional processing. This stage includes removing links, emoji, special characters, html entities, references for people, which are words starting with "@", punctuation marks, and numbers.

In the topic modeling operation, the pre-processed data receives new preparation operations which are the removal of hashtags and stopwords, which were removed with the NLTK

library.[4] The resulting texts were processed to remove inflections from words, using WordNetLemmatizer, also from the NLTK library, in English texts, and SpaCy[5] in Portuguese texts.

With these data, the topics were identified using the GSDMM algorithm. As presented in Section 3.1, in addition to the number of iterations, this method has three parameters to be configured: α, β and the maximum number of topics K. In their work, the authors presented a study on the influence of α and β on the topics produced. Using values close to those used by the authors, we performed empirical experiments to select these parameters. Thus, we considered α = 0.1 and β = 0.3 for the English dataset, and α = 0.1 and β = 0.1 for the Portuguese dataset.

After setting a maximum number of topics, this method is able to automatically determine the number of topics to be produced. However, for these datasets, the method produced a high number of topics with difficulty for semantically interpretation. Therefore, we chose to limit the number of topics, producing a set of more semantically complete topics. Thus, 30 topics were produced for English messages and 50 for Portuguese. Since in this technique, each message is related to only one topic, this information is added to the message record.

To choose the sentiment analysis technique that would be used in our data sets, we tested different data, attribute extraction techniques and classifiers. More details on these experiments can be found in Section 4.2. Analyzing the test results, we selected the CrystalFeel algorithm 2 for English messages. This method was applied to all messages, and the polarity and predominant emotion information was added to the message registers.

For Portuguese messages, a data preparation with the extraction of unigrams, bigrams and sentence embeddings was chosen. The method selected for the production of the embeddings was the Multilingual Universal Sentence Encoder for Semantic Retrieval, and the classifier was Logistic regression. The training used the same data and procedure indicated in Section 4.2. The result of sentiment analysis, which in this case was only polarity, was included in the message record.

The outputs of the natural language processing step are messages with the indication of topics, polarity and emotions, as represented on the right of Fig. 2. Another result is a vector with the participation of words in the topics. With these vectors, it is possible to extract, for example, the most important words for the topics. This was used in the analysis stage. Through the production of WordClouds and using the pyLDAvis[6] tool, it was possible to group similar topics, which corresponded to the most discussed subjects during the studied period.

Using the relationship between topics and subjects, and the relationship between messages and topics, it was possible to relate messages to subjects. These groups of messages were used to plot variations in volume and sentiment. Using variations in the volume of messages, we identified some periods of interest, which were used to search for related news. To do this, we used the Google search engine,[7] in the news tab, which allows us to configure a date range, language and location.

4.1. Dataset description

We collected tweets using the public streaming Twitter application programming interface (API) by filtering for general COVID-19-related keywords including: 'covid', 'corona', 'ncov', 'ncov-19', 'covid-19', 'pandemic' (*'pandemia'* in Portuguese), 'quarantine' (*'quarentena'* in Portuguese).

---

[3] https://nlp.stanford.edu/projects/snli/.

[4] https://www.nltk.org/.

[5] https://spacy.io.

[6] https://pypi.org/project/pyLDAvis/.

[7] https://www.google.com.
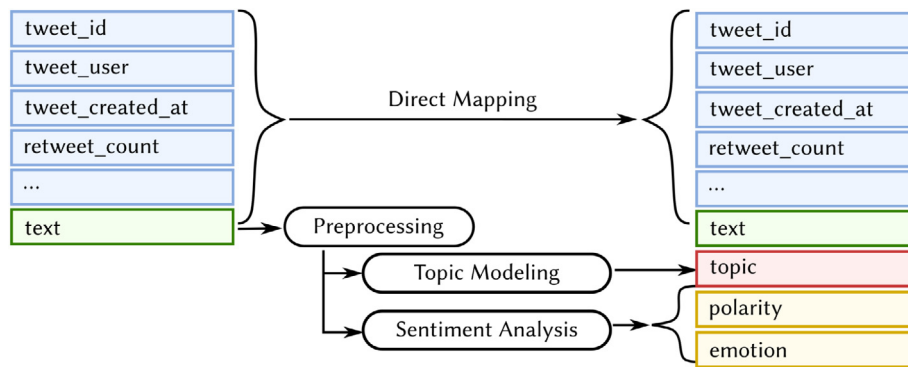
Fig. 2. Detailed Natural Language Processing step: the text from each tweet goes to a pre-processing (removing links, emoji, html, "@"), then, the tweets are employed into topic modeling and sentiment analysis algorithms, resulting in a set of topics and polarity/emotions classification for each tweet, respectively.

Table 1
SemEval 2018 - task 1. Emotion intensity regression (EI-reg).

	Pearson correlations (r)				
	Anger	Fear	Joy	Sadness	Avg.
CrystalFeel - word embedding	0.611	0.557	0.585	0.580	0.583
CrystalFeel	**0.702**	**0.689**	0.666	**0.689**	**0.686**
SBERT + Linear Regression	0.642	0.657	**0.697**	0.637	0.658
SBERT + Random Forest	0.587	0.615	0.651	0.636	0.622
SBERT + Linear SVM	0.642	0.653	0.574	0.638	0.626
SBERT + SVM	0.628	0.652	0.691	0.630	0.650

We extracted the COVID-19 tweets during the period between April 17, 2020 through August 08, 2020. The English dataset contains 7,144,349 tweets and the Portuguese dataset contains 7,125,530 tweets. After removing duplicate messages, the total remaining are 3,332,565 and 3,155,277, respectively.[8]

The geographic distribution of the messages was estimated with the users' addresses. Approximately 67% of the messages had the user's address, which was evaluated using the Geonames service.[9] For this task, we only considered addresses that had at least 2 alphanumeric characters.

This way, it was possible to identify the origin of 3,919,276 English messages, which corresponds to 54.9% of the total tweets retrieved in English. Considering the addresses, the most common countries were the USA (42.5%), India (10.8%), Canada (5.9%), and the United Kingdom (5.9%). For Portuguese tweets, we identified the origins of 3,735,963, which corresponds to 52.4% of the total tweets retrieved in Portuguese. The most common countries were Brazil (85%), Portugal (3.6%), and the United States (1.9%).

4.2. Sentiment analysis experiments

The sentiment analysis was evaluated using data from Se-mEval 2018 - task 1 [30]. These data are composed of messages collected from Twitter and manually labeled. The labels were assigned with Best–Worst Scaling (BWS) method, reducing inter- and intra-annotator inconsistencies. In this way, it was possible to build a data set that has not only a binary label, but its intensity, represented on a scale ranging from 0 (less intense) to 1 (more intense). Data is available in three languages, and organized into different learning tasks. In this work, we use the English data of emotion intensity (EI-reg) and valence/polarization intensity (V-reg).

Following the SemEval 2018 methodology, the results will be compared with Pearson correlations. The data set is provided with three partitions, which are the training, development and

test-gold data, which was used to rank the teams in the event's competition. Thus, to obtain the results presented in the following, the methods were calibrated only with the training data and evaluated on the development data. In all methods, we use the implementations of the Sklearn[10] library. The methods were executed with default parameters, with the exception of Logistic Regression, where the maximum number of interactions was increased to 2000.

The results indicated for CrystalFeel method were extracted from its original work. This work also contains the precision the authors found by applying only word embeddings. These values are indicated here as "CrystalFeel - word embedding". For the embeddings tests, we reduced to 50 dimensions using principal component analysis (PCA).

Table 1 shows the results for emotion intensity. CrystalFeel is a system that was developed specifically for Sentiment Analysis combining many features. This way, it presents the best results. On the other hand, Sentence BERT was developed for other applications and was trained with data different from those used in this experiment. Even so, the transfer of learning using the indicated classifiers, achieved reasonable results. Comparing with the CrystalFeel - word embedding, it is possible to see that SBERT is more advantageous in this problem.

Table 2 shows the results for valence/polarity regression. In addition to extracting attributes with SBERT, we also carry out experiments with Unigrams and Bigrams. To produce sentence embedding, the data was prepared in the same way as in the emotion data. To produce unigrams and bigrams, additionally, stop-words were removed and the words were changed to their canonical form (lemma) with NLTK tool. Rare words and bigrams, which would occur less than 3 times in the data set, have been removed.

Again, the best results were obtained by CrystalFeel. The approach of extracting attributes with SBERT and applying regression methods achieved close results, better than with unigrams and bigrams.

Table 2
SemEval 2018, task 1, Sentiment Scale Regression.

	Pearson correlations (r)
CrystalFeel	**0.816**
Unigram + bigram + Linear Regression	0.387
Unigram + bigram + Random Forest	0.496
Unigram + bigram + Linear SVM	0.388
Unigram + bigram + SVM	0.585[a]
SBERT + Linear Regression	0.764
SBERT + Random Forest	0.794
SBERT + Linear SVM	0.765
SBERT + SVM	0.761

[a]This value was calculated by the event team.

For Portuguese messages, we use a data set available on the Kaggle[11] platform. This data set was built using the same methodology applied in [31], where the training data was built with messages that were extracted from Twitter through a search for emoticons. The emoticons had been classified as positive or negative, and the messages that had these emoticons received the same classification, as a way to build a training set with noise labels.

This data set is divided into four parts. Two of them are messages classified as positive and negative, which differ in subject, with messages about politics in one and messages with no specific theme in another. There are also two files with messages extracted from news channels, which could be used as messages closer to neutral sentiment.

After evaluating the neutral messages, we noticed several noises, with messages that are closer to positive or negative classification, and, for this reason, we chose to discard this data. In addition, as our research is not related to politics, we also discard the file with this theme. In this way, we used only the file of positive and negative messages, without a defined theme, which was composed of 780,000 messages, with 33.5% positive and 66.5% negative.

Since works on sentiment analysis are less frequent in Portuguese than in English, we perform a larger volume of experiments, and these results are one of the contributions of this work. The experiments were organized in three groups, in the first we used only unigrams and bigrams, in the second only embeddings and in the third the combination of unigrams, bigrams and embeddings. All values indicated refer to the results obtained with a 10 cross-validation execution.

Table 3 shows the results considering six classifiers trained with unigrams and bigrams features. For a better understanding, the precision of positive and negative messages are represented separately. The classifiers that achieved highest results were Logistic Regression, Random Forest and Linear SVM. The positive set has small F1-score compared to negative, probably because it is the minority class.

In the following tests we only consider the classifiers that achieved the highest results in Table 3. Table 4 shows the results of the classification with sentence embedding. For fastText we use the pre-trained networks provided by the Interinstitutional Center for Computational Linguistics [32]. To convert word embeddings into sentence embeddings, the same function as fastText was applied, which is the average of the vectors normalized by l2-norm. For SBERT, we use the pre-trained network "xlm-r-bert-base-nli-stsb-mean-tokens",[12] for mUSE, we use the version provided in this same package "distiluse-base-multilingual-cased-v2", both networks of the multilingual package. The results

---

[11] https://www.kaggle.com/augustop/portuguese-tweets-for-sentiment-analysis.

[12] https://www.sbert.net/docs/pretrained_models.html.

Table 3
Portuguese sentiment classification considering unigrams and bigrams as features.

Classifier	Negative		Positive	
	Precision	F1-score	Precision	F1-score
Naive Bayes	**0.83**	0.85	0.72	0.63
Logistic Regression	0.81	**0.86**	0.77	0.66
Random Forest	0.81	**0.86**	**0.78**	0.65
Linear SVM	0.81	**0.86**	0.77	**0.66**
MLP	0.82	0.83	0.66	0.65
AdaBoost	0.75	0.83	0.76	0.51

Table 4
Portuguese sentiment classification considering sentence embedding as features.

Classifier	Negative		Positive	
	Precision	F1-score	Precision	F1-score
fastText - 50 attributes				
Logistic Regression	0.73	0.81	0.65	0.44
Random Forest	0.75	0.83	0.73	0.51
Linear SVM	0.72	0.81	0.66	0.41
fastText - 300 attributes				
Logistic Regression	0.79	**0.84**	0.73	0.60
Random Forest	0.76	**0.84**	0.77	0.54
Linear SVM	0.78	**0.84**	0.74	0.49
Multilingual Universal Sentence Encoder				
Logistic Regression	0.79	**0.84**	0.71	0.61
Random Forest	0.77	**0.84**	**0.75**	0.57
Linear SVM	0.79	**0.84**	0.72	0.61
Sentence BERT				
Logistic Regression	**0.80**	**0.84**	0.71	**0.63**
Random Forest	0.78	**0.84**	0.72	0.59
Linear SVM	**0.80**	**0.84**	0.71	0.62

Table 5
Portuguese sentiment classification combining unigrams, bigrams and sentence embedding as features.

Classifier	Negative		Positive	
	Precision	F1-score	Precision	F1-score
Unigrams, Bigrams and fastText 50 attributes				
Logistic Regression	0.82	0.86	0.77	0.68
Linear SVM	0.82	0.86	0.76	0.68
Unigrams, Bigrams and fastText 300 attributes				
Logistic Regression	0.83	**0.87**	0.77	0.70
Linear SVM	0.83	0.86	0.76	0.69
Unigrams, Bigrams and mUSE				
Logistic Regression	0.84	**0.87**	**0.78**	**0.71**
Linear SVM	0.83	**0.87**	0.77	0.70
Unigrams, Bigrams and SBERT				
Logistic Regression	0.80	0.84	0.71	0.63
Linear SVM	0.80	0.84	0.72	0.62

achieved by Logistic Regression, Random Forest and Linear SVM were similar. All the sentence embeddings performed similarly too, except fastText with 50 attributes.

Table 5 shows the results combining unigrams, bigrams and sentence embeddings. With this combination we achieved the best results, with emphasis on the unigrams formed with mUSE, which achieved the best results for positive and negative messages.

5. Results

Following, Section 5.1 presents the main topics selected from English and Portuguese tweets. Section 5.2 describes the variations in message volumes by topic, moreover, we associate the

Table 6
English topics.

Topic	Correlated words
Economic impacts	Work, impact, business, crisis, pay
Case reports/statistics	Case, death, report, die, number, people, total, patient, record, update, confirm
Proliferation care	Mask, wear, people, school, test, social, spread, reopen
Politics	Trump, government, response, president, fauci(Anthony Stephen Fauci), state, minister
Entertainment	Watch, like, video, love, show
Treatments	Vaccine, patient, hospital, drug, trial, treatment, plasma, hydroxychloroquine
Online events	Join, webinar, live, talk, discuss, impact, virtual, tomorrow, host
Charity	Support, help, fund, relief, donate, provide, community, food, million
Sports	Player, season, football, team, league, game, play, sport
Anti-racism protests	Protest, police, kill, black, american, death, die, right, war

peaks with news reported during the period, in order to identify some facts that may have motivated these variations. Section 5.3 presents the sentiment analysis over the retrieved tweets.

5.1. Topics emerged from tweets

In this section, the topics identified for the two languages will be presented. Using the results of topic modeling (complete list in Appendix), we analyze the most relevant terms, Word-Clouds produced in the messages related to the topics, and some sample messages, to relate topics to subjects. Table 6 presents the English topics and the most correlated terms. The subjects ranged from directly related to COVID-19, such as treatments, proliferation care and case reports, to indirectly related subjects, such as economy, education, sports and politics.

Other subjects emerged from the situation imposed such as charity and online events. Moreover, we identified messages about anti-racism protests that were not motivated by the coronavirus epidemic, but they occurred in the analyzed period.

Table 7 presents the Portuguese topics and the most correlated terms. The words listed here have been translated into English for a general understanding. The topics also ranged from subjects directly related to COVID-19, such as treatments, proliferation, case reports, to subjects indirectly related, such as economy, education, sports and politics. A different subject that appeared in Brazilian Portuguese was everyday life.

Fig. 3 represents the volume of messages in the identified topics. The horizontal axis represents the percentage of the dataset assigned to the topics. The topics are ordered, starting from the most frequent. It is possible to notice that similar topics received different attention in the two datasets.

In Fig. 2, we observed several similarities between the topics found, which were proliferation care, case reports and statistics, economic impacts, politics, treatments, entertainment and sports. The differences in volumes may have been influenced by the political moment since we are in an election year in the USA.

Anti-racism protests are not directly related to the covid epidemic but were collected with keywords related to the epidemic. After analyzing some of the messages, we observed some reports about the quarantine interruption during the protests.

The case statistics are produced mainly by news channels and, therefore, they are expected to be among the subjects of the greatest repercussion. In addition to this, we note that, in both languages, messages about economic impacts and proliferation care are among the most discussed issues.

In Portuguese dataset, we observed the subject of health and beauty, with several reports on anxiety caused by confinement and variations in body weight. In addition to these, some messages were found about changes in appearance, such as new haircuts, as a form of entertainment during the quarantine.

5.2. Results of volume analysis

This section shows the variations in daily message volumes for the topic. In Figs. 4 and 5 the horizontal axis represents the posting dates and the vertical axis the number of posts. In these figures, the blue lines represent the total messages. The other lines are related to sentiment analysis and will be discussed in the next section. To simplify the visualization, the numbers represent the sum of the messages over a week.

Fig. 4 shows the variations in message volumes by topic for English data. In this figure, we can see some moments when there was a considerable increase in the volume of posts. Using the keywords indicated in Table 6 and the dates with the highest volume of messages, we conducted a search for news in order to identify some facts that may have motivated these variations.

From the beginning of our data series in April to the end of May, we noticed a large volume of messages related to economics and politics. In this period, in Europe, leaders approve support measures for companies and workers [33]. In addition, small companies received loans to help them to return to business [34]. In the USA an additional payment is discussed for essential workers, such as grocery [35]. In the same period, there were also many messages related to politics. In our search for related news, we find President Donald Trump's justification for a controversial comment [36]. We also find news that the president has indicated his intention to remove the USA from the list of WHO members [37]. Dr. Anthony Stephen Fauci was also quoted several times, as in the warning about the importance of preventing cities from resuming their activities too quickly [38], and also indicating that there is no scientific evidence that the coronavirus appeared in a Chinese laboratory [39].

In the same period, we observed a large volume of messages related to treatments and proliferation care. Among the news about treatment, we found different alternatives under study. One is Vaccine, which received $8 Billion in donations from world leaders [40] and which has already started testing on humans in the UK [41]. We also found much news about the drug Remdesivir that, in clinical tests, reduced the length of hospital stay for advanced patients [42]. In addition to these, we find the Food and Drug Administration (FDA) alerts on the risks of *hydroxychloroquine* [43], and WHO stops testing with this drug [44]. Regarding proliferation care, we found much news about the president's refusal to use masks [45]. In addition, some regions of the USA are preparing to reopen [46]. At the end of May, more than 200,000 deaths [47] are reported worldwide, 50,000 in the USA [48].

At the beginning of June, it is possible to notice a reduction in the volume of messages for all selected subjects, which coincides with the anti-racism protests, that was important not only in the USA [49], but worldwide [50]. That same month, we observed the return of message volumes for the selected subjects and, in treatments, researchers from England say that there is evidence that the drug *dexamethasone* can reduce the deaths on severely

Table 7
Portuguese topics.

Topic	Correlated words
Economic impacts	Buying, money, working, crisis
Treatments	Vaccine, cure, chloroquine, virus, treatment, patient, hydroxychloroquine, medical
Proliferation care	To stick, house, to leave, people, to pass, mask, a party
Case reports/statistics	Death, case, number, dead, register, confirm, mother, country, father uncle
Education and culture	Class, learn, how to, study, read, school, teacher, college
Sports	Playing, football, player, Flamengo(a football club), team, club, championship
Politics	Bolsonaro, governor, president, stf(Supreme Federal Court), minister, major, approve, law, project, mp(Public Prosecutor's Office or provisional measure), chamber, federal, public
Heath and beauty	Hair, fattening, health, mental, painting, losing weight, anxiety
Entertainment	Watch, series, musician, movie, listen, season, lauch, netflix, show, album, live, clip
Daily life	Day, sleep, wake up, eat, photo, play, night, morning
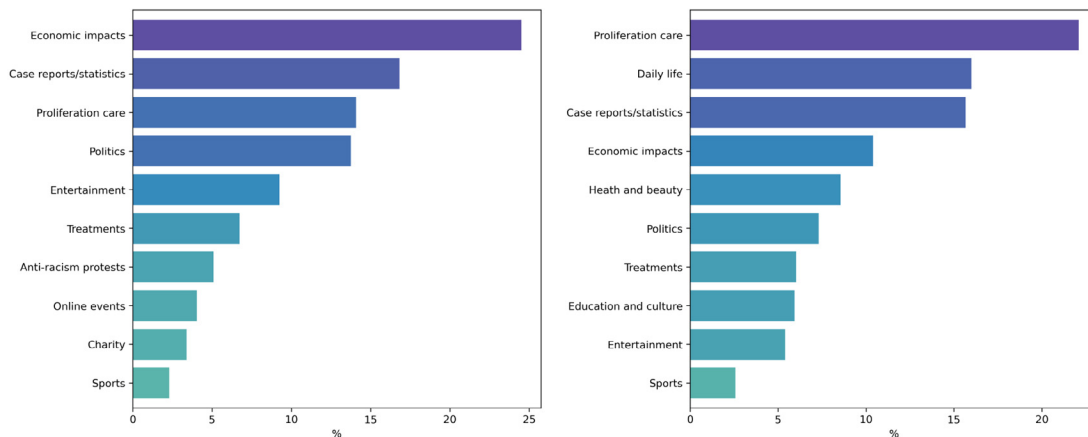
Fig. 3. Volume of messages by topic (English data on the left and Portuguese data on the right).
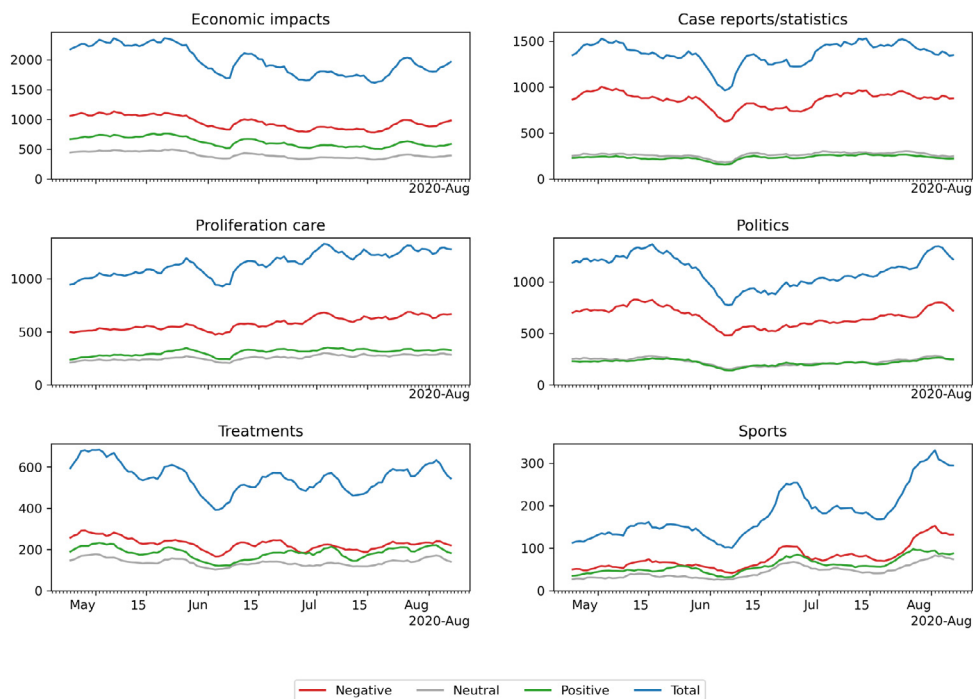
Fig. 4. English volume variations for topic. The horizontal axis represents the posting dates where each point represents the sum of the messages over a week and the vertical axis the number of posts. The blue lines represent the total messages. The other lines are related to sentiment analysis where red are negative and green are positive.

Fig. 5. Portuguese volume variations for topic. The horizontal axis represents the posting dates where each point represents the sum of the messages over a week and the vertical axis the number of posts. The blue lines represent the total messages. The other lines are related to sentiment analysis where red are negative and green are positive.

ill patients [51]. Studies indicate that the use of masks can re-duce the transmission [52], and some states start to require this item [53].

At this time, the USA was officially in an economic crisis caused by the coronavirus [54]. In proliferation care, we found some news about the reopening of schools [55], and about diffi-culties with hybrid school schedules [56]. We also found a lot of news about wearing masks, which highlight its importance [57], regions that started to demand them [58], and people's adher-ence [59]. There was also a discussion about the possibility of transmission by people who are infected but have no symp-toms [60]. The number of victims reaches 400,000 worldwide, and the UK reaches 50,000 dead. In the USA, hospitalizations reach records [61].

At the end of June, we noticed an increase in the volume of sports-related messages. The Premier League football games [62] are back, and Major League Baseball also plans to return, but first needs to overcome some difficulties [63].

In July, we noticed an increase in the volume of messages related to statistics. At the beginning of that period, the news indicated a record number of cases in the USA [64], mainly in Florida [65]. Also during this month, there were record cases in India, which became the third most affected country [66] and in Australia, even with the declaration of the lockdown [67].

Towards the end of June, in economics, there were several wage changes. Some companies have stopped paying additional [68], or reduced wages [69]. In proliferation care, Donald Trump started wearing masks [70], and encouraged Americans to do

the same [71]. There were also some large stores that started to require their customers to wear masks [72]. Plans to reopen schools continue, but the return is postponed in many states, while classes remain remote [73]. In treatments, authorities re-port that Russian [74] and Chinese [75] hackers are trying to steal data from vaccine development research. Despite this, tests continue, and the results are promising [76]. In this same period, we observed an increase in the volume of messages related to sports, in the same period Major League Baseball restarts [77]. Unlike football competitions, which remain paralyzed [78].

Fig. 5 shows the volume variations by subject for Portuguese data. In this figure, we can see some moments when there was a considerable increase in the volume of posts. In the same way that it was done for the English data, we use the keywords of Table 7, and the dates with greater variations in the number of messages, to search for news that may have motivated these variations.

In chronological order, some subjects start the series with a large volume of news. Among them are economics and politics. During this period, one of the news with the greatest impact was the payment of emergency aid, which was initially paid to 6 million Brazilians [79]. Considering that most of the aid amounts will be withdrawn immediately, the central bank starts to mon-itor the system, fearing scarcity of notes [80]. We also observed some news describing the impacts of the crisis [81,82]. In politics, the health minister changes [83]. The president of the republic expands the list of essential services, which is allowed to operate during the quarantine [84]. In addition, the president made some

controversial statements, which had great repercussions [85,86]. There have been some changes in legislation to help tackle the pandemic [87,88].

Also at the beginning of the series, we found some news related to the proliferation care, some news described parties that were happening, despite the recommendation of isolation [89]. We also found news about decrees that require wearing masks in public places [90] and the reaction of the population [91]. About education, some families stopped sending children to school [92]. In order to prevent classes from being interrupted, some distance learning options have been proposed [93,94]. However, distance learning has some obstacles such as teacher preparation [95] and the lack of student resources, such as internet access [96]. During this period, there was a significant increase in cases in the Northeast of the country [97] and Brazil exceeded 10,000 deaths [98].

Treatment-related news in mid-May indicates promising initial results for the vaccine under development by the University of Oxford [99]. During this period, there were several discussions about the efficacy of *hydroxychloroquine* in the treatment of COVID-19, motivated by studies that indicated that this drug was ineffective for the disease [100], and the interruption of tests carried out at WHO [101]. Despite this, the government continued to recommend using it [102] even in mild conditions [103].

At the end of May and the beginning of June, we noticed an increase in the volume of messages on proliferation care, economics and politics. The news from this period reported that hospitals have reached high levels of occupancy [104] and that new beds are being made available [105]. With the increase in the number of cases, the USA prohibits the entry of Brazilians [106]. After a month in office, the health minister, Nelson Teich, resigns from the position for disagreeing with the president's opinion on the use of *hydroxychloroquine* [107]. In economics, the fourth payment of emergency aid is confirmed [108]. The Chamber of Deputies approves a provisional measure that allows the reduction of hours and wages during the pandemic. This measure still needs Senate approval [109].

In June, we saw a significant increase in the volume of messages related to politics. During this period, there was a change in the publication of epidemic data, hiding total infections and deaths [110]. After the reaction of parties opposed to the government, the supreme federal court determined that the data should be fully disclosed again [111]. Also in June, the government of Sao Paulo announced the agreement between the Chinese laboratory Sinovac and the Butantan[13] institute, to carry out tests with the CoronaVac vaccine [112]. In addition, we also observed an important variation in the volume of messages related to sports, perhaps motivated by the resumption of football championships [113].

Also starting in June, but extending to July, we noticed a large volume of messages about proliferation care. During this period, the Federal Court, through an injunction, determined that the president should wear masks in public spaces in the Federal District [114]. The Attorney General Office appeals against this obligation and manages to overturn the decision [115,116]. Bolsonaro sanctions the law that makes the use of masks in public spaces mandatory [117], with vetoes for the requirement in commerce, schools and temples [118]. On July 7, the president announces that he has COVID-19 [119].

Also between June and July, there was a large volume of messages related to education. During this period, we found news related to the preparations for the return of schools [120]. We also find news related to the challenge of distance education during quarantine [121].

In the first half of July, there was an increase in messages related to the economy. In this period, we find news about the impact of the crisis on business [122], and also on problems involving the payment of emergency aid [123]. In this period, Brazil is the country with the highest daily number of cases and deaths [124] and takes second place in the total number of cases [125].

Throughout July until the end of the series, we observed an increase in the volume of messages related to treatments. During this period, tests with the coronavirus vaccine begin in Sao Paulo [126]. The Oxford vaccine has shown promising results in clinical studies [127]. Companies and the Lemann[14] foundation support the national manufacture of the vaccine [128]. At the end of July and beginning of August, the planning for the return to school continues, that should occur by dividing the classes so that some students continue in distance learning to reduce crowding in classrooms [129]. In this period, Brazil is approaching 100,000 deaths, with more than 1000 cases per day [130].

5.3. Results of sentiment analysis

Figs. 4 and 5 show the polarities of the messages. In red are the messages with negative polarity and in green the positive ones. The values indicated on the vertical axis refer to the message count classified in the different polarities. To simplify the visualization, the numbers represent the sum of the messages over a week.

As expected, since they are data about the pandemic, most messages are negative. However, looking at the figures, we realize that in some topics this difference is more significant. The topic with the highest percentage of negative messages, in both languages, was proliferation care, with 63% of messages in English and 60% of messages in Portuguese. The topic with the highest volume of positive messages in English were treatments, totaling 33% of messages. In Portuguese, treatment was the third most positive topic, with 63% of messages. In Portuguese messages, the topic with the highest volume of positive messages was politics, with approximately 75% of messages classified as positive.

For English texts, a more detailed analysis of the feelings was possible, using CrystalFeel. In this technique, most positive messages are classified as joy, and negative messages are divided into anger, sadness and fear. Considering only the negative messages, which are the most frequent, Fig. 6 represents the proportion of messages in different feelings. In this figure, anger and fear were the most common.

6. Discussion

The key topics we identified were representative of the public conversations being had in news outlets during April and August. We notice some news stimulate a spike especially in the topics 'Economic impacts' (2000 tweets per week) and 'Case reports/statistics' (1500 tweets per week) for English data and 'Proliferation care' (4000 tweets per week) and 'Case reports/statistics' (2500 tweets per week) for Portuguese data.

Negative emotions were dominant during the COVID-19 pandemic for almost all the topics in English. In Portuguese negative was prevalent in 'Proliferation care', 'Case reports and statistics' and 'Education and culture'. For 'Economic impacts' the sentiments are almost equivalent. Given the prevalence of negative sentiments, it is important a strategic public health communication and actions to maintain the public's mental well-being during the pandemic.

---

13 http://www.butantan.gov.br/.
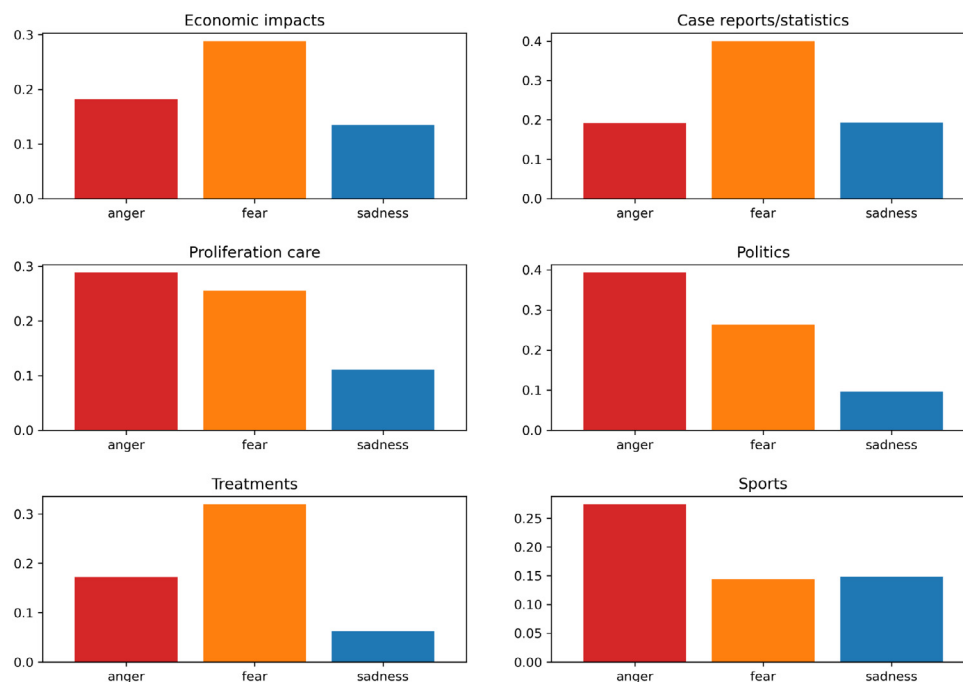
14 https://fundacaolemann.org.br/.

Fig. 6. Proportion of emotions related to anger, fear and sadness for the English topics.

Our topic findings and sentiment analysis differ from other works mentioned on Related Work that analyze COVID-19 and Twitter. In [16] authors analyzed Twitter data at the beginning of the pandemic (February 2 to March 15) and from 12 identified topics, 10 of them had positive sentiment and only 2 had negative. Authors in [20] analyzed data from January 01 to March 23 and also found a high number of positive and neutral tweets. Only the re-tweets are most negative. Probably, in the beginning of the pandemic, people were optimistic and thinking about the problems caused by the virus will be solved quickly. However, as time passed and the problems remained the negative sentiments started to be predominant.

Other studies, such as [131], which did a review of the psychological impact of quarantine confirmed negative psychological impacts, such as post-traumatic stress symptoms, confusion, and anger. The paper cites the following stressors factors: "longer quarantine duration, infection fears, frustration, boredom, inadequate supplies, inadequate information, financial loss, and stigma". Our results comply with that article since fear is the highest sentiment regarding case reports/statistics, treatments, and economic impacts, and anger is the highest sentiment regarding proliferation care and politics in English data.

A challenge we find in this work was to perform sentiment analysis for Portuguese, since few annotated datasets are available to train the algorithms. We have a gap in the literature regarding the study and development of new techniques for processing languages other than English.

Overall, sentiments about COVID-19 are rapidly evolving and the issues surrounding the pandemic are challenging and complex. In this situation, effective information/communication is very important. Public health officials should make clear the importance of quarantine, trustful case reports and treatments should be announced, to avoid frustration/stress, misinterpreted information and fake news diffusion in the population.

7. Conclusion

Social networks have become a popular tool not only for advertising but for idea dissemination and individual opinion-making. Analyzing social network contents can give us a perception of society and the world. In the current situation caused by COVID-19, understanding the emotions of the people are extremely important.

In this paper, we explored Twitter content in English and Portuguese mainly from the USA and Brazil in response to the COVID-19 pandemic from April through August 2020. We found ten main topics related to COVID-19 in both languages, where seven topics are equivalent.

Negative emotions were dominant in almost all the topics identified during the COVID-19 pandemic. Most of them are related to proliferation care, case reports and statistics. This pattern was similar for English and Portuguese tweets analyzed. These negative emotions are expected given the worldwide level of this pandemic. These sentiments could be counterbalanced by governments and authorities' strategic public health communication.

One limitation of this work is the keywords employed to retrieve content related to COVID-19. It is possible that some relevant tweets were missed if they did not include the keywords. Our study only included tweets from Brazil and the USA, on the basis that these countries had a large number of COVID-19 cases. Further research could evaluate the usage of Twitter by different countries.

Furthermore, the findings reported in this study are limited to only those that use Twitter. Therefore, caution is advised before assuming the generalizability of the results, as Twitter is not used by everyone in the population.

Future works could explore different algorithms and data analysis in Portuguese as well as other less spoken languages. Studies comparing behavioral changes, emotions and impacts surrounding the COVID-19 pandemic with different countries are welcome too.

CRediT authorship contribution statement

Klaifer Garcia: Software, Validation, Data curation, Methodology, Visualization, Writing - original draft. **Lilian Berton:** Conceptualization, Methodology, Visualization, Writing - original draft, Writing - review & editing, Supervision.

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.

Appendix

List of topics by subject for english data

Economic impacts

Topic 1: new, work, help, need, health, business, support, time, impact, change. Topic 2: health, people, crisis, world, coronavirus, many, need, impact, face, country, economic, global, time, community, care. Topic 3: pay, get, people, work, due, job, worker, money, lose, go, million, need, make, help, business. Topic 4: coronavirus, market, economy, economic, impact, global, new, say, due, business, industry, crisis, year, amid, via, world, company, hit, job, report

Case reports/statistics

Topic 1: death, people, case, rate, die, new, virus, study, flu, coronavirus. Topic 2: case, new, death, report, total, number, record, day, update, confirm, coronavirus, today, test, positive, state. Topic 3: die, hospital, patient, family, one, home, people, get, year, day, say, nurse, lose, old, friend

Politics

Topic 1: trump, say, president, coronavirus, response, fauci, test, via, rally, house, white, donald. Topic 2: government, say, uk, health, minister, response, coronavirus, johnson, public, boris, make, death, handle, call, china. Topic 3: trump, state, bill, vote, say, relief, court, federal, new, coronavirus, house, health, government, senate, election

Treatments

Topic 1: vaccine, drug, trial, test, use, treatment, say, study, patient, coronavirus, hydroxychloroquine, cure, virus, first, new. Topic 2: patient, hospital, test, need, plasma, contact, trace, app, recover, bed, care, help, use, get, blood

Sports

Topic 1: test, player, positive, season, football, due, team, league, game, play, sport, say, get, coronavirus, go

List of topics by subject for portuguese data

Proliferation care

Topic 1: furar, fazer, ir, ver, pessoa, ficar, gente, todo, postar, sair, achar, amigo, casa, festa. Topic 2: ir, querer, ficar, acabar, casa, furar, sair. Topic 3: saude, combater, teste, hospital, governar, novo, sobre, estar, durante, fazer, prefeitura. Topic 4: caso, estar, cidade, dia, leito, hospital, sp, ir, rio, novo, morte, aumentar, número, comércio, paulo, uti, prefeito, ver, isolamento. Topic 5: máscara, ir, usar, pessoa, fazer, poder, casa, ficar, ter, todo, vírus, sair, mão, passar, álcool

Case reports/statistics

Topic 1: ir, meu, casa, mãe, fazer, pai, dia, ficar, morrer, pegar, pessoa, tio, todo, trabalhar, hoje, aqui, família. Topic 2: brasil, morte, mil, morto, número, dia, pessoa, caso, país, morrer, ir, milhão. Topic 3: caso, morte, brasil, novo, registrar, confirmar, mil, óbito, número, hora, último, coronavírus, dia, chegar, estar, total, morto, boletim, país. Topic 4: ir, ficar, dor, achar, dia, pegar, ter, saber, sintoma, sentir, todo, falto, ansiedade, fazer, cabeça, ar, medo, morrer, pensar, crise. Topic 5: morrer, ano, após, hospital, vítima, internar, dia, positivar, paciente, ver, dizer, médico

Economic impacts

Topic 1: brasil, país, dizer, novo, coronavírus, crise, ver, poder, governar, dever, mundo, estar, economia, brasileiro. Topic 2:fazer, ir, ajudar, trabalhar, poder, gente, nessa, ter, contar, querer, durante, saber, dar, todo, dinheiro, casa, dever, meio, pagar. Topic 3: ir, comprar, usar, nessa, acabar, fazer, roupas, querer, dia, sair, ficar, coisa, casa, dinheiro, todo, ter, gastar

Politics

Topic 1: bolsonaro, governador, governar, fazer, ir, presidente, saude, estar, stf, ministro, poder, prefeito, combater, querer, brasil. Topic 2: durante, bolsonaro, governar, aprovar, saude, lei, combater, medir, projeto, stf, sobre, mp, contra, estar, câmara, público, federal, deputar, presidente

Treatments

Topic 1: vacinar, ir, curar, tomar, pro, querer, contra, vírus, achar, dia, acordar, descobrio. Topic 2: cloroquina, usar, tratamento, contra, hidroxicloroquina, médico, dizer, curar, paciente, remédio, sobre, estudar, ver, tomar, fazer, poder, ivermectina, tratar. Topic 3: vacinar, contra, poder, dizer, testar, estudar, brasil, teste, coronavírus, ver, novo, vírus, pesquisar, oms

Education and culture

Topic 1: ir, aula, fazer, ter, ano, voltar, querer, escola, dia, todo, estudar, acabar, dar, passar, meio, trabalhar, professor, saber, faculdade. Topic 2: fazer, nessa, ir, coisa, dia, aprender, ler, querer, começar, todo, saber, ter, estudar, algum, achar, conseguir.

Sports

Topic 1: voltar, jogar, ir, futebol, jogador, fazer, flamengo, testar, ver, time, clube, todo, meio, ter, campeonato.

References

[1] R.T. Gandhi, J.B. Lynch, C. del Rio, Mild or moderate Covid-19, New Engl. J. Med. (2020) http://dx.doi.org/10.1056/NEJMcp2009249.

[2] S. L, B. AM, P. K, M. C, U. L, M. RM., Twitter as a tool for health research: A systematic review, Am J Public Health 107 (1) (2017).

[3] S. Jordan, S. Hovet, I..-H. Fung, H. Liang, K.-W. Fu, Z. Tse, Using Twitter for public health surveillance from monitoring and prediction to public response, Data 4 (6) (2019).

[4] S.R. Rufai, C. Bunce, World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis, J. Publ. Health (2020) http://dx.doi.org/10.1093/pubmed/fdaa049.

[5] Twitter. Quarterly results: 2019 fourth quarter, 2020, URL https://investor.twitterinc.com/financial-information/quarterly-results/default.aspx. (Accessed: 2020-08-02).

[6] Twitter. Testing promoted tweets on our logged-out experience, 2015, URL https://blog.twitter.com/en_us/a/2015/testing-promoted-tweets-on-our-logged-out-experience.html. (Accessed: 2020-08-02).

[7] M. Thelwall, S. Thelwall, Retweeting for COVID -19: Consensus building, information sharing, dissent, and lockdown life, 2020, arXiv:2004.02793.

[8] E.M. Glowacki, A.J. Lazard, G.B. Wilcox, M. Mackert, J.M. Bernhardt, Identifying the public's concerns and the Centers for Disease Control and Prevention's reactions during a health crisis: An analysis of a Zika live Twitter chat, Am. J. Infect. Control 44 (12) (2016) 1709–1711, http://dx.doi.org/10.1016/j.ajic.2016.05.025.

[9] A. Stefanidis, E. Vraga, G. Lamprianidis, J. Radzikowski, P.L. Delamater, K.H. Jacobsen, D. Pfoser, A. Croitoru, A. Crooks, Zika in Twitter: Temporal variations of locations, actors, and concepts, JMIR Publ. Health Surveill 3 (2) (2017) 1709–1711, http://dx.doi.org/10.2196/publichealth.6925.

[10] C. Morin, I. Bost, A. Mercier, J.-P. Dozon, L. Atlani-Duault, Information circulation in times of Ebola: Twitter and the sexual transmission of ebola by survivors, PLoS Curr. 1 (10) (2018) http://dx.doi.org/10.1371/currents. outbreaks.4e35a9446b89c1b46f8308099840d48f.

[11] M. Miyabe, A. Miura, E. Aramaki, Use trend analysis of Twitter after the great east Japan earthquake, in: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion, Association for Computing Machinery, 2012, pp. 175–178.

[12] K. Lachlan, Z. Xu, E. Hutter, R. Adam, P. Spence, A little goes a long way: Serial transmission of Twitter content associated with hurricane irma and implications for crisis communication, J. Strateg. Innov. Sustain. 14 (1) (2019) http://dx.doi.org/10.33423/jsis.v14i1.984.

[13] Z. Mottaghinia, M.-R. Feizi-Derakhshi, L. Farzinvash, P. Salehpour, A review of approaches for topic detection in Twitter, J. Exp. Theoret. Artif. Intell. (2020) 1–27, http://dx.doi.org/10.1080/0952813X.2020.1785019.

[14] H. Rosa, J.P. Carvalho, F. Batista, Detecting a tweet's topic within a large number of Portuguese Twitter trends, in: 3rd Symposium on Languages, Applications and Technologies, in: OpenAccess Series in Informatics (OASIcs), vol. 38, 2014, pp. 185–199.

[15] V.K. Rangarajan Sridhar, Unsupervised topic modeling for short texts using distributed representations of words, in: Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing, Association for Computational Linguistics, Denver, Colorado, 2015, pp. 192–200, http://dx.doi.org/10.3115/v1/W15-1526.

[16] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, Z. Shah, Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study, J. Med. Internet Res. 22 (4) (2020) e19016, http://dx.doi.org/10.2196/19016.

[17] T. Mackey, V. Purushothaman, J. Li, N. Shah, M. Nali, C. Cardier, B. Liang, M. Cai, R. Cuomo, Machine learning to detect self-reporting of symptoms, testing access, and recovery associated with COVID-19 on Twitter: Retrospective big data infoveillance study, JMIR Publ. Health Surveillance 6 (2) (2020) http://dx.doi.org/10.2196/19509.

[18] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C.M. Valensise, E. Brugnoli, A.L. Schmidt, F.Z. Paola Zola, A. Scala, The COVID-19 social media infodemic, 2020, arXiv:2003.05004.

[19] M.O. Lwin, J. Lu, A. Sheldenkar, P.J. Schulz, W. Shin, R. Gupta, Y. Yang, Global sentiments surrounding the COVID-19 pandemic on Twitter: Analysis of Twitter trends, JMIR Publ. Health Surveill 6 (2) (2020) e19447, http://dx.doi.org/10.2196/19447.

[20] K. Chakraborty, S. Bhatia, S. Bhattacharyya, J. Platos, R. Bag, A.E. Hassanien, Sentiment analysis of COVID-19 tweets by deep learning classifiers— A study to show how popularity is affecting accuracy in social media, Appl. Soft Comput. 97 (2020) 106754, http://dx.doi.org/10.1016/j.asoc. 2020.106754.

[21] R.K. Gupta, A. Vishwanath, Y. Yang, COVID-19 Twitter dataset with latent topics, sentiments and emotions attributes, 2020, arXiv:2007.06954.

[22] A. Alamoodi, B. Zaidan, A. Zaidan, O. Albahri, K. Mohammed, R. Malik, E. Almahdi, M. Chyad, Z. Tareq, A. Albahri, H. Hameed, M. Alaa, Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review, Expert Syst. Appl. (2020) 114155, http://dx.doi.org/ 10.1016/j.eswa.2020.114155.

[23] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent Dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.

[24] J. Qiang, Z. Qian, Y. Li, Y. Yuan, X. Wu, Short text topic modeling techniques, applications, and performance: A survey, IEEE Trans. Knowl. Data Eng. 14 (8) (2020) 1, http://dx.doi.org/10.1109/tkde.2020.2992485, arXiv:1904.07695.

[25] J. Yin, J. Wang, A Dirichlet multinomial mixture model-based approach for short text clustering, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, pp. 233–242, http://dx.doi.org/10.1145/2623330.2623715.

[26] R.K. Gupta, Y. Yang, Crystalfeel at semeval-2018 task 1: Understanding and detecting emotion intensity using affective lexicons, in: Proceedings of the 12th International Workshop on Semantic Evaluation, Association for Computational Linguistics, New Orleans, Louisiana, 2018, pp. 256–263, http://dx.doi.org/10.18653/v1/s18-1038.

[27] Y. Yang, D. Cer, A. Ahmad, M. Guo, J. Law, N. Constant, G.H. Abrego, S. Yuan, C. Tar, Y.-H. Sung, B. Strope, R. Kurzweil, Multilingual universal sentence encoder for semantic retrieval, 2019, arXiv:1907.04307.

[28] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, 2019, arXiv preprint arXiv:1908.10084.

[29] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, 2018, arXiv preprint arXiv:1810.04805.

[30] S.M. Mohammad, F. Bravo-Marquez, M. Salameh, S. Kiritchenko, SemEval-2018 task 1: Affect in tweets, in: Proceedings of International Workshop on Semantic Evaluation (SemEval-2018), New Orleans, LA, USA, 2018.

[31] A. Go, R. Bhayani, L. Huang, Twitter Sentiment classification using distant supervision, Processing (2009) 1–6, URL https://cs.stanford.edu/people/ alecmgo/papers/TwitterDistantSupervision09.pdf.

[32] N. Hartmann, E. Fonseca, C. Shulby, M. Treviso, J. Rodrigues, S. Aluisio, Portuguese word embeddings: Evaluating on word analogies and natural language tasks, 2017, arXiv:1708.06025.

[33] Eurogroup statement on the pandemic crisis support, Consilium (2020) URL https://www.consilium.europa.eu/en/press/press-releases/2020/05/ 08/eurogroup-statement-on-the-pandemic-crisis-support/.

[34] H. Treasury, Small businesses boosted by bounce back loans, 2020, URL https://www.gov.uk/government/news/small-businesses-boosted-by-bounce-back-loans.

[35] S. Selasky, Grocery Employees Say They Work in Fear, Want Hazard Pay to Continue, Detroit Free Press, 2020, URL https: //www.freep.com/story/news/local/michigan/2020/05/21/kroger-grocery-store-essential-workers-hazard-pay/5227345002/.

[36] K.S. Allyson Chiu, Trump claims controversial comment about injecting disinfectants was 'sarcastic', 2020, URL https://www.washingtonpost.com/ nation/2020/04/24/disinfectant-injection-coronavirus-trump/,

[37] L. Egan, Trump says U.S. will be 'terminating' relationship with WHO, 2020, URL https://www.nbcnews.com/politics/white-house/trump-says-u-s-will-be-terminating-relationship-who-n1218441,

[38] J. Herb, L. Fox, Fauci tells congress that states face serious consequences if they reopen too quickly, 2020, URL https://www.cnn.com/2020/05/12/ politics/anthony-fauci-congress-hearing/index.html,

[39] P.b.A. Harnik, P.b.A. Wong, P.b.G. Tames, Fauci: No scientific evidence the coronavirus was made in a chinese lab, 2020, URL https://www.nationalgeographic.com/science/2020/05/anthony-fauci-no-scientific-evidence-the-coronavirus-was-made-in-a-chinese-lab-cvd/,

[40] M. Stevis-gridneff, L. Jakes, World leaders join to pledge $8 billion for vaccine as U.S. goes it alone, N.Y. Times (2020) URL https://www.nytimes. com/2020/05/04/world/europe/eu-coronavirus-vaccine.html,

[41] S. Dean, Covid-19 Vaccine trial on humans starts as UK warns restrictions could stay in place until next year, 2020, URL https://www.cnn.com/2020/ 04/23/health/coronavirus-vaccine-trial-uk-gbr-intl/index.html,

[42] nih clinical trial shows remdesivir accelerates recovery from advanced covid-19_2020, 2020, URL https://www.nih.gov/news-events/news-releases/nih-clinical-trial-shows-remdesivir-accelerates-recovery-advanced-covid-19.

[43] C.R. Laurie McGinley, FDA Warns about hydroxychloroquine dangers, citing serious heart issues, including death, 2020, URL https://www.washingtonpost.com/health/2020/04/24/fda-warns-about-hydroxychloroquine-dangers-citing-serious-heart-issues-including-death/,

[44] S.a. agencies, WHO Halts hydroxychloroquine trial for coronavirus amid safety fears, 2020, URL https://www.theguardian.com/world/2020/may/ 25/who-world-health-organization-hydroxychloroquine-trial-trump-coronavirus-safety-fears,

[45] C. Cillizza, People *really* want donald trump to wear a mask in public, 2020, URL https://edition.cnn.com/2020/05/21/politics/donald-trump-mask-polling/index.html,

[46] P. Svitek, Gov. Greg abbott to let restaurants, movie theaters and malls open with limited capacity friday, 2020, URL https://www.texastribune. org/2020/04/27/texas-reopening-coronavirus-greg-abbott/,

[47] Coronavirus: Worldwide death toll climbs to 200,000, 2020, URL https: //www.bbc.com/news/world-52428162,

[48] Coronavirus: US death toll passes 50,000 in world's deadliest outbreak, 2020, URL https://www.bbc.com/news/world-us-canada-52415579,

[49] George floyd death: Violence erupts on sixth day of protests, 2020, URL https://www.bbc.com/news/world-us-canada-52872401,

[50] A. Sandford, George Floyd: Protests continue in Europe despite virus restrictions, 2020, URL https://www.euronews.com/2020/06/06/black-lives-matter-protesters-take-to-streets-in-europe-despite-pandemic-restrictions,

[51] M. Marchione, Cheap drug is first shown to improve COVID-19 survival, 2020, URL https://apnews.com/article/ 89d963958b042cc921e64ab3eff5a74d,

[52] A. Kim, Best way to reduce coronavirus transmission is by wearing a face mask, study finds, 2020, URL https://www.cnn.com/2020/06/12/health/ coronavirus-mask-wellness-trnd/index.html,

[53] A. Kim, S. Andrew, J. Froio, These are the states requiring people to wear masks when out in public, 2020, URL https://www.cnn.com/2020/06/19/ us/states-face-mask-coronavirus-trnd/index.html,

[54] J. Frankel, The US is officially in recession thanks to the coronavirus crisis | Jeffrey Frankel, 2020, URL https://www.theguardian.com/business/2020/ jun/16/us-recession-coronavirus-crisis,

[55] J. Wisely, Gov. Whitmer says schools will reopen this fall for in-person learning, 2020, URL https://www.freep.com/story/news/education/2020/ 06/17/gov-whitmer-schools-reopen-fall-person-learning/3205842001/,

[56] D.R. Superville, Hybrid school schedules: More flexibility; big logistical challenges, 2020, URL https://www.edweek.org/ew/articles/2020/06/25/hybrid-school-schedules-more-flexibility-big-logistical.html.

[57] K. Kelland, Widespread mask-wearing could prevent COVID-19 second waves: study, 2020, URL https://www.reuters.com/article/us-health-coronavirus-masks-study-idUSKBN23G37V,

[58] California orders people wear masks indoor spaces, 2020, URL https://abcnews.go.com/Health/wireStory/california-orders-people-wear-masks-indoor-spaces-71327757,

[59] K.M. Terry Tang, US Virus outbreaks stir clash over masks, personal freedom, 2020, URL https://apnews.com/54374ff841dfd84323a1fb86d1e93180,

[60] N.H.-D. Will Feuer, Asymptomatic spread of coronavirus is 'very rare,' WHO says, 2020, URL https://www.cnbc.com/2020/06/08/asymptomatic-coronavirus-patients-arent-spreading-new-infections-who-says.html,

[61] C.B. Samantha Pell, Coronavirus hospitalizations rise sharply in several states following memorial day, 2020, URL https://www.washingtonpost.com/health/2020/06/09/coronavirus-hospitalizations-rising/,

[62] B. Church, English premier league returns after 100-day hiatus, 2020, URL https://edition.cnn.com/2020/06/17/football/premier-league-return-coronavirus-spt-intl/index.html,

[63] ABC News, ABC News Network, 2020, URL https://abcnews.go.com/Sports/mlb-commissioner-rob-manfred-now-confident-2020-season/story?id=71264260.

[64] A. Watts, US Sees a record number of new coronavirus cases reported in a single day, 2020, URL https://www.cnn.com/2020/07/01/health/us-coronavirus-wednesday/index.html,

[65] L. Shumaker, Florida sets one-day record with over 15,000 new COVID cases, more than most countries, 2020, URL https://www.reuters.com/article/us-health-coronavirus-usa-idUSKCN24D0HE,

[66] H. Sullivan, Global report: India becomes third worst-affected country as giant Covid-19 hospital opens, 2020, URL https://www.theguardian.com/world/2020/jul/06/global-report-india-sees-record-daily-rise-as-capital-opens-giant-covid-19-hospital,

[67] Coronavirus: Australia's victoria records huge case jump, 2020, URL https://www.bbc.com/news/world-australia-53589817,

[68] N. Meyersohn, Retailers are canceling coronavirus hazard pay. That's a mistake, 2020, URL https://www.cnn.com/2020/07/20/business/hazard-pay-essential-workers-coronavirus/index.html,

[69] Saurabh Sinha / TNN / Updated: Jul 23, 2020, Air India salary cut: Air India pilots warn of 'potentially disastrous psychological impact': India business news - times of India, 2020, URL https://timesofindia.indiatimes.com/business/india-business/pay-cut-air-india-pilots-warn-of-potentially-disastrous-psychological-impact/articleshow/77129881.cms,

[70] J. Diamond, K. Liptak, 'Pleading' from aides led to trump agreeing – after months – to wear a mask, 2020, URL https://www.cnn.com/2020/07/10/politics/trump-mask-walter-reed-coronavirus/index.html,

[71] G. Segers, Trump encourages Americans to wear masks, and warns pandemic may "get worse before it gets better", 2020, URL https://www.cbsnews.com/news/coronavirus-trump-wear-face-mask-pandemic-get-worse/,

[72] N. Meyersohn, Walmart, kroger will start requiring customers in US stores to wear masks, 2020, URL https://www.cnn.com/2020/07/15/business/walmart-masks/index.html,

[73] A. Grayer, Several big US school districts are extending remote classes into the fall, 2020, URL https://www.cnn.com/2020/07/13/us/school-reopening-plans-major-cities/index.html,

[74] O.b.D.K. Sepkowitz, Opinion: Russia, the US and the Covid-19 vaccine free-for-all, 2020, URL https://edition.cnn.com/2020/07/19/opinions/vaccine-covid-19-russia-hack-us-sepkowitz/index.html,

[75] J.E. Barnes, U.S. Accuses hackers of trying to steal coronavirus vaccine data for China, N.Y. Times (2020) URL https://www.nytimes.com/2020/07/21/us/politics/china-hacking-coronavirus-vaccine.html,

[76] N. Bar-Zeev, W.J. Moss, Encouraging results from phase 1/2 COVID-19 vaccine trials, Lancet 396 (10249) (2020) 448–449, http://dx.doi.org/10.1016/s0140-6736(20)31611-1.

[77] T. Goldman, Baseball shouts 'play ball!' but for how long? A nervous 2020 season begins, 2020, URL https://www.npr.org/2020/07/23/894461036/baseball-shouts-play-ball-but-for-how-long-a-nervous-2020-season-begins,

[78] K. Bonagura, Big Sky moves football season to the spring, wants FCS conferences to follow, 2020, URL https://www.espn.com/college-football/story/_/id/29611581/big-sky-conference-planning-football-games-spring,

[79] A.O. Globo, Auxílio emergencial: Caixa aguarda autorização do governo para pagar a 6 milhões de pessoas, 2020, URL https://revistapegn.globo.com/Noticias/noticia/2020/05/auxilio-emergencial-caixa-aguarda-autorizacao-do-governo-para-pagar-6-milhoes-de-pessoas.html,

[80] M. Barbosa, Banco central monitora sistema temendo falta de dinheiro em bancos, 2020, URL https://www.correiobraziliense.com.br/app/noticia/economia/2020/05/10/internas_economia,853364/banco-central-monitora-sistema-temendo-falta-de-dinheiro-em-bancos.shtml,

[81] R. Hessel, Crise do coronavírus também pode derrubar PIB brasileiro de 2021, 2020, URL https://www.correiobraziliense.com.br/app/noticia/economia/2020/05/11/internas_economia,853393/crise-do-coronavirus-tambem-pode-derrubar-pib-brasileiro-de-2021.shtml,

[82] Sul e Sudeste sofrem mais com crise da pandemia que outras regiões, 2020, URL https://economia.ig.com.br/2020-04-30/sul-e-sudeste-sofrem-mais-com-crise-da-pandemia-que-outras-regioes.html,

[83] T. Molina, Nelson Teich, o comandante escolhido pelo governo contra o coronavírus, 2020, URL https://veja.abril.com.br/podcast/nelson-teich-o-comandante-escolhido-pelo-governo-contra-o-coronavirus/,

[84] Bolsonaro amplia lista de atividades consideradas essenciais na pandemia do coronavírus, 2020, URL https://g1.globo.com/politica/noticia/2020/04/29/bolsonaro-amplia-lista-de-atividades-consideradas-essenciais-na-pandemia-do-coronavirus.ghtml,

[85] 'E daí? Lamento. Quer que eu faça o quê?', diz Bolsonaro sobre mortes por coronavírus; 'Sou Messias, mas não faço milagre', 2020, URL https://g1.globo.com/politica/noticia/2020/04/28/e-dai-lamento-quer-que-eu-faca-o-que-diz-bolsonaro-sobre-mortes-por-coronavirus-no-brasil.ghtml,

[86] Bolsonaro diz que cobrança sobre mortes por coronavírus precisa ser feita a governadores e prefeitos, 2020, URL https://g1.globo.com/politica/noticia/2020/04/29/bolsonaro-diz-que-cobranca-sobre-mortes-por-coronavirus-precisa-ser-feita-a-governadores-e-prefeitos.ghtml,

[87] Coronavírus: Lei torna uso de máscara obrigatório em locais públicos de todo o Paraná, 2020, URL https://diariodotransporte.com.br/2020/04/28/coronavirus-lei-torna-uso-de-mascara-obrigatorio-em-locais-publicos-de-todo-o-parana/,

[88] Senado aprova projeto que cria linha de crédito para micro e pequenas empresas, 2020, URL https://g1.globo.com/politica/noticia/2020/04/24/senado-aprova-projeto-que-cria-linha-de-credito-para-micro-e-pequenas-empresas.ghtml,

[89] GCM Interrompe festa com mais de 40 pessoas durante a quarentena em campo limpo Paulista, 2020, URL https://g1.globo.com/sp/sorocaba-jundiai/noticia/2020/04/20/gcm-interrompe-festa-com-mais-de-40-pessoas-durante-a-quarentena-em-campo-limpo-paulista.ghtml,

[90] Uso de máscara em locais públicos é obrigatório em todo o Paraná, 2020, http://www.aen.pr.gov.br/modules/noticias/article.php?storyid=106752%5C,

[91] Folhapress, Mulher morre em confusão por máscara em hipermercado no Paraná, 2020, URL https://gauchazh.clicrbs.com.br/seguranca/noticia/2020/04/mulher-morre-em-confusao-por-mascara-em-hipermercado-no-parana-ck9kewet100o201o4xtvqn91t.html,

[92] Famílias decidem tirar crianças da escola por causa da pandemia, 2020, URL https://cbn.globoradio.globo.com/media/audio/302369/familias-decidem-tirar-criancas-da-escola-por-caus.htm,

[93] C. Moreira, Aula on-line valerá como presença na rede pública do DF, 2020, URL https://www.correiobraziliense.com.br/app/noticia/cidades/2020/05/20/interna_cidadesdf,856932/aula-on-line-valera-como-presenca-na-rede-publica-do-df.shtml,

[94] Conexão Escola MG: como acessar as aulas online da rede estadual, 2020, URL https://www.techtudo.com.br/listas/2020/05/conexao-escola-mg-como-acessar-as-aulas-online-da-rede-estadual.ghtml,

[95] R. Cafardo, 8 em cada 10 professores não se sentem aptos a aulas online, 2020, URL https://www.terra.com.br/noticias/educacao/8-em-cada-10-professores-nao-se-sentem-aptos-a-aulas-online,d293f995588bfc1c448ddb3cbf5bc3e2j3ntiyx6.html,

[96] Sem internet, merenda e lugar para estudar: veja obstáculos do ensino à distância na rede pública durante a pandemia de Covid-19, 2020, URL https://g1.globo.com/educacao/noticia/2020/05/05/sem-internet-merenda-e-lugar-para-estudar-veja-obstaculos-do-ensino-a-distancia-na-rede-publica-durante-a-pandemia-de-covid-19.ghtml,

[97] Bahia registra mais três mortes de pacientes com coronavírus e número sobe para 100; total de casos é de 2.676, 2020, URL https://g1.globo.com/ba/bahia/noticia/2020/04/29/bahia-registra-mais-tres-mortes-de-pacientes-com-coronavirus-e-numero-sobe-para-100-numero-total-de-casos-e-de-2676.ghtml,

[98] Brasil entra na lista dos 6 países que ultrapassaram a barreira dos 10 mil mortos por Covid-19; veja comparativo, 2020, URL https://g1.globo.com/bemestar/coronavirus/noticia/2020/05/09/brasil-entra-na-lista-dos-6-paises-que-ultrapassaram-a-barreira-dos-10-mil-mortos-por-covid-19-veja-comparativo.ghtml,

[99] R. Galileu, Vacina contra Covid-19 testada nos EUA tem resultados iniciais promissores, 2020, URL https://revistagalileu.globo.com/Ciencia/noticia/2020/05/vacina-contra-covid-19-testada-nos-eua-tem-resultados-iniciais-promissores.html,

[100] Estudo com hidroxicloroquina não constata melhora de pacientes hospitalizados com Covid-19, 2020, URL https://g1.globo.com/bemestar/coronavirus/noticia/2020/05/08/estudo-com-pacientes-hospitalizados-em-nova-york-nao-constata-melhora-no-tratamento-com-hidroxicloroquina-para-covid-19.ghtml,

[101] C. Barreto, OMS suspende o uso da cloroquina e hidroxicloroquina em testes contra a Covid-19, 2020, URL https://pebmed.com.br/oms-suspende-o-uso-da-cloroquina-e-hidroxicloroquina-em-testes-contra-a-covid-19/,

[102] Bolsonaro 'exige' que ministro da Saúde recomende a cloroquina, 2020, URL https://valor.globo.com/politica/noticia/2020/05/14/bolsonaro-exige-que-ministro-da-sade-recomende-a-cloroquina.ghtml,

[103] Ministério mantém recomendação para uso de cloroquina em casos leves de Covid-19, mas corrige texto, 2020, URL https://g1.globo.com/bemestar/coronavirus/noticia/2020/05/21/ministerio-da-saude-divulga-versao-de-orientacoes-sobre-cloroquina-com-asinaturas.ghtml,

[104] Santa Casa de Votuporanga tem 90% de ocupação nos leitos de UTI para Covid-19, 2020, URL https://g1.globo.com/sp/sao-jose-do-rio-preto-aracatuba/noticia/2020/06/01/santa-casa-de-votuporanga-tem-90percent-de-ocupacao-nos-leitos-de-uti-para-covid-19.ghtml,

[105] Coronavírus: Santa Casa de Mogi Mirim inaugura seis novos leitos de UTI para Covid-19, 2020, URL https://g1.globo.com/sp/campinas-regiao/noticia/2020/06/01/coronavirus-santa-casa-de-mogi-mirim-inaugura-seis-novos-leitos-de-uti-para-covid-19.ghtml,

[106] Casa Branca cita avanço da Covid-19 para justificar proibição à entrada de brasileiros nos EUA - Mundo - Diário do Nordeste, 2020, URL https://diariodonordeste.verdesmares.com.br/mundo/casa-branca-cita-avanco-da-covid-19-para-justificar-proibicao-a-entrada-de-brasileiros-nos-eua-1.2249103,

[107] Teich sobre saída do Ministério da Saúde: 'Não houve alinhamento com o presidente', 2020, URL https://www.em.com.br/app/noticia/politica/2020/05/24/interna_politica,1150288/teich-sobre-saida-do-ministerio-nao-houve-alinhamento-com-bolsonaro.shtml,

[108] E.d. Minas, Bolsonaro confirma quarta parcela do auxílio emergencial: valor deve cair, 2020, URL https://www.correiobraziliense.com.br/app/noticia/economia/2020/05/29/internas_economia,859308/bolsonaro-confirma-quarta-parcela-do-auxilio-emergencial.shtml,

[109] Câmara conclui aprovação de MP que reduz jornada e salários durante pandemia, 2020, URL https://valor.globo.com/politica/noticia/2020/05/28/camara-conclui-aprovacao-de-mp-que-reduz-jornada-e-salarios-durante-pandemia.ghtml,

[110] J.S.e.M. Vargas, Mudança de divulgação ocorreu após Bolsonaro exigir número de mortes por covid abaixo de mil por dia - Saúde, 2020, URL https://saude.estadao.com.br/noticias/geral,mudanca-de-divulgacao-ocorreu-apos-bolsonaro-exigir-numero-de-mortes-abaixo-de-mil-por-dia,70003328525,

[111] J. Oliveira, Moraes do STF manda Governo Bolsonaro retomar divulgação total de dados da covid-19, 2020, URL https://brasil.elpais.com/brasil/2020-06-09/governo-bolsonaro-aprofunda-confusao-sobre-a-covid-19-com-a-publicacao-de-dados-incompletos.html,

[112] Doria anuncia que Butantan será parceiro de laboratório chinês para vacina contra o coronavírus em fase final de testes, 2020, URL https://g1.globo.com/sp/sao-paulo/noticia/2020/06/11/governo-de-sp-diz-que-instituto-butantan-vai-produzir-vacina-contra-o-coronavirus.ghtml,

[113] Campeonato Carioca pode voltar com jogo do Flamengo já no próximo final de semana, 2020, URL https://diariodorio.com/campeonato-carioca-pode-voltar-com-jogo-do-flamengo-ja-no-proximo-final-de-semana/,

[114] Justiça Federal obriga Bolsonaro a usar máscara em espaços públicos do DF, 2020, URL https://g1.globo.com/df/distrito-federal/noticia/2020/06/23/justica-federal-obriga-bolsonaro-a-usar-mascara-em-espacos-publicos-do-df.ghtml,

[115] AGU recorre da decisão que obriga Bolsonaro a usar máscara quando sair em público em Brasília, 2020, URL https://g1.globo.com/df/distrito-federal/noticia/2020/06/26/agu-recorre-da-decisao-que-obriga-bolsonaro-a-usar-mascara-quando-sair-em-publico-em-brasilia.ghtml,

[116] Justiça derruba decisão que obrigava Bolsonaro a usar máscara em locais públicos do DF, 2020, URL https://g1.globo.com/df/distrito-federal/noticia/2020/06/30/justica-derruba-decisao-que-obrigava-bolsonaro-a-usar-mascara-em-locais-publicos-do-df.ghtml,

[117] Bolsonaro sanciona lei que torna obrigatório o uso de máscara, 2020, URL https://agenciabrasil.ebc.com.br/saude/noticia/2020-07/bolsonaro-sanciona-lei-que-torna-obrigatorio-o-uso-de-mascara,

[118] D. Gullino, Bolsonaro veta uso obrigatório de máscara em comércio, escolas e templos, 2020, URL https://oglobo.globo.com/sociedade/bolsonaro-veta-uso-obrigatorio-de-mascara-em-comercio-escolas-templos-1-24513437,

[119] Bolsonaro informa em rede social que novo exame para a Covid-19 deu resultado positivo, 2020, URL https://g1.globo.com/politica/noticia/2020/07/15/bolsonaro-anuncia-em-rede-social-que-novo-exame-para-a-covid-19-deu-resultado-positivo.ghtml,

[120] Protocolo prevê turmas divididas e quarentena em escola que tiver caso de Covid-19 no Paraná, 2020, URL https://g1.globo.com/pr/parana/educacao/noticia/2020/07/16/protocolo-preve-turmas-divididas-e-quarentena-em-escola-que-tiver-caso-de-covid-19-no-parana.ghtml,

[121] Com digitalização de salas de aula, pandemia acentua exclusão escolar, 2020, URL https://agenciabrasil.ebc.com.br/educacao/noticia/2020-07/com-digitalizacao-de-salas-de-aula-pandemia-acentua-exclusao-escolar,

[122] Pandemia está se transformando em uma crise de empregos muito pior que a de 2008, diz OCDE, 2020, URL https://g1.globo.com/economia/noticia/2020/07/07/pandemia-esta-se-transformando-em-uma-crise-de-empregos-muito-pior-que-a-de-2008-diz-ocde.ghtml,

[123] A. Gavioli, Dinheiro do auxílio emergencial desaparece da conta de correntistas, 2020, URL https://www.infomoney.com.br/minhas-financas/dinheiro-do-auxilio-emergencial-desaparece-da-conta-de-correntistas/,

[124] M.E.C. Bruna Lima, Covid-19: Brasil é o país que registra atualmente mais casos e mortes, 2020, URL https://www.correiobraziliense.com.br/app/noticia/brasil/2020/06/11/interna-brasil,862912/covid-19-brasil-e-o-pais-que-registra-atualmente-mais-casos-e-mortes.shtml,

[125] Brasil passa Reino Unido e se torna o 2% país com mais mortes por Covid-19, 2020, URL https://www1.folha.uol.com.br/equilibrioesaude/2020/06/brasil-passa-reino-unido-e-se-torna-o-2o-pais-com-mais-mortes-por-covid-19.shtml,

[126] Teste de vacina contra coronavírus começa em SP, e médica é 1ª voluntária, 2020, URL https://www.uol.com.br/vivabem/noticias/redacao/2020/07/21/teste-de-vacina-contra-coronavirus-comeca-em-sp-e-medica-e-1-voluntaria.htm,

[127] M. Ansede, Vacina de Oxford contra o coronavírus oferece resultados promissores em teste com 1.000 pessoas, 2020, URL https://brasil.elpais.com/ciencia/2020-07-20/vacina-de-oxford-contra-o-coronavirus-oferece-resultados-promissores-em-teste-com-1000-pessoas.html,

[128] Fundação Lemann e empresas investirão R$100 mi em fábrica de vacina contra Covid-19 no Brasil, 2020, URL https://epocanegocios.globo.com/Empresa/noticia/2020/08/epoca-negocios-fundacao-lemann-e-empresas-investirao-r100-mi-em-fabrica-de-vacina-contra-covid-19-no-brasil.html,

[129] B.A.e.Y. Setubal, Metade na escola, metade em casa: ensino híbrido é o próximo nó da educação em tempos de pandemia, 2020, URL https://oglobo.globo.com/sociedade/educacao/metade-na-escola-metade-em-casa-ensino-hibrido-o-proximo-no-da-educacao-em-tempos-de-pandemia-24554153,

[130] Com quase 100 mil mortes, Brasil registra 1.058 óbitos em 24 horas, 2020, URL https://g1.globo.com/bemestar/coronavirus/noticia/2020/08/07/casos-e-mortes-por-coronavirus-no-brasil-em-7-de-agosto-segundo-consorcio-de-veiculos-de-imprensa.ghtml,

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