# Portuguese emotion detection model using BERTimbau applied to COVID-19 news and replies

Francisco Bráulio Oliveira[1][0000−0003−2926−8214] and Jaime Simão Sichman[1][0000−0001−8924−9643]

Laboratório de Técnicas Inteligentes (LTI)
Escola Politécnica (EP)
Universidade de São Paulo (USP), Brazil
{fcobraulio@usp.br, jaime.sichman}@usp.br

**Abstract.** In this study, we developed an Emotion Detection (ED) model for Brazilian Portuguese social media comments and used it to analyze emotion distribution in replies to COVID-19-related news on Twitter. We translated the GoEmotions dataset into Portuguese using a Large Language Model, preserving the original social media writing style, and fine-tuned the BERTimbau model for detecting Ekman's six basic emotions plus neutral emotion. We collected tweets related to the COVID-19 pandemic from major Brazilian news portals and analyzed user replies for expressed emotions. Chi-square tests indicated that emotion distribution depended on both the news topic and the media outlet, while the Augmented Dickey-Fuller (ADF) test showed that some emotions' distributions were influenced by the publication date. Among replies with identifiable emotions, anger (21.1%), joy (17.2%), and surprise (11.6%) were most prevalent. The Brazilian pandemic inquiry (CPI) prompted the highest anger prevalence (41.9%), while news from UOL Notícias had the highest anger (40.0%) and g1 led in sadness (13.6%).

**Keywords:** NLP · Emotion Detection · Deep Learning · Topic Modeling · Covid-19 · News.

## 1  Introduction

Emotions constitute a fundamental aspect of human cognition and behavior. They play a critical indicator of psychological state and social interaction. Understanding the emotional expressions within text not only provides insights into individual sentiment, but also unveils collective trends and behavioral patterns within communities. Therefore, analysis of textual data using Artificial Intelligence (AI) techniques has became an important activity across multiple domains. In recent years, the intersection of natural language processing (NLP) and machine learning has facilitated significant advancements in emotion analysis. Notably, the utilization of deep learning architectures, such as BERT (Bidirectional Encoder Representations from Transformers), achieve significant results.

Social media platforms provide a significant amount of data to be analyzed by text classification models. Also, sentiments and topics on social media can be a good proxy for public opinion. For instance, data from social media, such as Twitter, replicate consumer confidence and Presidential job approval polls [28] and contribute to understand the response to terrorist attacks [6]. Recently, it helps elucidate COVID-19-related discourse, providing valuable insights into public sentiment, reactions to health policies, and the dissemination of information during crises.

However, while numerous studies have explored emotion analysis in English-language contexts, research in Brazilian Portuguese remains limited. This gap in the literature restricts comprehensive analyses of emotions expressed by Portuguese-speaking communities, particularly in response to significant events like the COVID-19 pandemic. Addressing this gap is crucial for promoting inclusive research and understanding the unique cultural nuances present in Brazilian Portuguese discourse.

In this study, we aim to contribute to the advancement of Natural Language Processing (NLP) research in Brazilian Portuguese by developing and applying a BERT-based emotion classification model, using the BERTimbau language model. Specifically, we seek to answer the following research questions:

RQ1: Given an Ekman-based labeled emotion corpus in English, is it feasible to use LLMs to generate a similarly labeled corpus in Brazilian Portuguese while maintaining the language style? Moreover, can we build an efficient emotion transformer-based classifier in Brazilian Portuguese using this new corpus?

RQ2: What are the predominant emotions expressed in comments on Brazilian media news about the COVID-19 pandemic on Twitter?

RQ3: How do these emotions distribute across different press outlets, news themes, and over time?

To address these questions, we employ a multi-step methodology involving corpus translation, model fine-tuning, dataset collection, and emotion detection. This study also aims to provide insights for policymakers, healthcare experts, and marketers by explaining the emotional landscape of COVID-19 discourse in Brazilian Portuguese. It hopes to enable more informed decision-making and communication strategies in the context of a major event. The remainder of the paper is organized as follows. Section 2 gives an overview of the related work. Section 3 describes our methodology, and Section 4 presents our findings. We conclude the paper in Section 5 with some general remarks.

## 2    Background and Related Work

Human emotions are expressed in a multi-modal fashion, encompassing facial expressions, voice, and language [23]. However, classification strategies focusing only on textual data exhibit performance comparable to multimodal approaches [1]. A crucial aspect in developing emotion classifiers is the establishment of an emotion model, defining how emotions are represented. The main kinds of emotion models include discrete and dimensional models [46]. Discrete models

categorize emotions into a limited set of categories. Paul Ekman's study suggests six fundamental emotions — anger, sadness, disgust, joy, surprise, and fear — originating from distinct neural systems [11]. In contrast, Robert Plutchik's model suggests that emotions occur in opposing pairs at varying intensities, augmenting Ekman's set with acceptance/trust and anticipation [32]. Meanwhile, dimensional models propose emotions as non-independent entities, represented within a continuous multi-dimensional space. Plutchik's two-dimensional wheel [32] and Russell's circumplex model [34] differentiate emotions based on arousal (activation vs. deactivation) and valence (pleasantness vs. unpleasantness). Furthermore, the Pleasure-Arousal-Dominance (PAD) model delineates emotions across three axes [22].

The field of research known as Affective Computing emerged with the recognition that computational systems could potentially classify a user's emotional state [30], from where emerged the Text-Based Emotional Detection (TBED). Annotated collections of documents, or texts, with distinct emotions facilitated the development of artificial intelligence models capable of supervised emotion detection in new documents, as well as performance evaluation and comparison with other models. The first annotated corpus for emotions was the ISEAR dataset, encompassing 7 distinct emotions [35]. Subsequently, Wordnet was developed as one of the primary lexical databases specifically constructed for Natural Language Processing (NLP) in English [12].

Lexical databases consist of structured sets of terms aiming to approximate the terms native speakers associate with a given concept. Early NLP applications focused on emotion detection used lexical databases to identify the predominant emotion in a document [38, 13, 40, 45]. Following this, word2vec [25] enabled the usage of word embeddings, opening the way for neural networks and deep learning in a variety of NLP tasks, including emotion detection. In 2017, the introduction of Transformers [42] enabled the development of language models, including BERT [10], representing a qualitative leap in emotion analysis tasks. Presently, TBED research indicates that lexicon-based applications are still in use; however, classifiers based on machine learning and deep learning are becoming more prevalent [2, 16, 15]. Recent works have employed Large Language Models (LLMs) for TBED activities [3, 44].

Among works focused on the Portuguese language, notable efforts include the development of WordnetAffectBR, one of the earliest and most significant lexical databases in Brazilian Portuguese [29]. Another contribution constructed a corpus of short news items (headlines and brief descriptions) and utilized Latent Semantic Analysis to classify documents among Ekman's six emotions, along with neutral emotion. More recently, a BERT-based language model named BERTimbau was developed [37]. BERTimbau underwent fine-tuning using a version of the GoEmotions corpus [9] translated by Google Translate to generate a model specialized in emotion classification in Portuguese [14]. This model was utilized to evaluate the polarity of comments related to political discourse, and it was compared to ChatGPT, demonstrating superior performance [36]. In contrast, other techniques, such as classification models based on lexical databases, were

employed to identify emotions in tweets related to the COVID-19 pandemic in Portuguese [4]. Regarding tweets related to the COVID-19 pandemic written in English, multiple studies used machine learning and deep learning techniques to investigate the distribution of emotions among social media users, finding a prevalence of negative emotions such as anger [47, 18, 27].

## 3   Methodology

In this study, we construct a model for classifying emotions in social media comments, applied to a dataset of COVID-19-related tweets collected from Twitter. We adapt the GoEmotions dataset [9] to Portuguese using ChatGPT to ensure that we keep the writing style of social media comments. Then, we fine-tune a pre-trained Portuguese language model, BERTimbau, for Emotion Detection (ED), employing Ekman's model of six basic emotions plus neutral emotion. Following this, we gather tweets related to the COVID-19 pandemic from major Brazilian news portals on Twitter, and their user replies. The fine-tuned model is used to detect the emotions expressed in these replies. Finally, we conduct statistical analysis on the distribution of these emotions relating them to the news topic, the news media, and the date when the news was published.

### 3.1   Model training

**Table 1.** Sample comments from GoEmotions, their translation to Portuguese using Google Translate, and chatGPT (GoEmotions-BR).

| GoEmotions [9] | Google Translate | GoEmotions-BR |
|---|---|---|
| I brought my love a cher-rrryyyyy.... | Trouxe uma cereja para meu amor.... | Eu trouxe umaaaa cerreeee-jjjaaaaa pro meu amoooor-rrrrr.... |
| *tomorrow Thank you for this. | *amanhã Obrigado por isso. | amanhã vlw por isso. |
| Thats super fricking cool I love it | Isso é super legal, eu adorei | Isso é muito legal pra caramba, adoro! |
| Lucky, when I did it my nickname ended up being [NAME]. | Sorte, quando fiz isso meu apelido acabou sendo [NOME]. | Sortudo, quando fiz isso, meu apelido acabou sendo [NOME]. |
| Yeah until the last one has creepily wed and repro-duced. | Sim, até que o último se case e se reproduza assustadora-mente. | Sim, até que o último tenha se casado de forma assusta-dora e reproduzido. |
| Looks better imo | Parece melhor, imo | Parece melhor na minha opinião. |

**Language Model (LM)** Language Models are artifacts capable of perform-ing language-related tasks automatically. Typically, they are generated through

machine learning or statistical modeling. Recently, LMs have been build using Deep Neural Network (DNN) models trained on large volumes of data, utilizing a transformer architecture [42]. One of the most commonly used LM architectures is the Bidirectional Encoder Representations from Transformers (BERT) [10]. This modeling approach provides contextual language features, enabling LMs to have generalist capabilities. That is, they can undergo fine-tuning processes, generating new models specialized in specific tasks, including sentiment analysis, machine translation, question-answering systems, and more. In this study, we utilize a LM called BERTimbau [37], trained using the BERT architecture on a dataset collected in Brazilian Portuguese. BERTimbau was trained on the brWaC corpus [43], compiled from web pages, consisting of 3.53 million documents and a total of 2.68 billion tokens.

**Fine-tuning** Fine-tuning a language model (LM) typically involves retraining one of its layers using an annotated dataset tailored to a specific task. This study aimed to develop an LM specialized in Emotion Detection (ED) using the GoEmotions dataset, which includes about 58,000 Reddit comments manually annotated into 27 emotions, later aggregated into Ekman's six basic emotions plus neutral. Hammes and Freitas [14] previously translated this dataset using Google Translate but noted its inability to preserve the social media writing style.

Recent studies [8, 17] have used ChatGPT for translating corpora to fine-tune LMs for underrepresented languages. Following this approach, we created GoEmotions-BR by translating the GoEmotions corpus with chatGPT-3.5-turbo, maintaining the original writing style and localizing slang (see Table **??**).

The translation prompt used was: "Translate the following message to Portuguese. Preserve the writing style, using equivalent slang and abbreviations, if needed." For example, "Looks better imo" was accurately translated to "Parece melhor na minha opinião," with chatGPT recognizing "imo" as "na minha opinião". Similarly, "I brought my love a cherrrryyyyy" was translated as "Eu trouxe umaaaa cerreeeejjjaaaaa pro meu amoooorrrrrr", preserving the original's expressive style, unlike Google Translate.

The GoEmotions-BR corpus was then used to fine-tune the BERTimbau-base model, creating BERTimbau-Ekman-BR, specialized in ED. We followed the code and configurations provided by Hammes and Freitas [14].

### 3.2 Emotion distribution

**The corpus** We utilize the Reuters Institute's Digital News Report 2023 [26] to identify the most prominent digital media outlets in Brazil. Then, we search for the highly active Twitter accounts that best represent each media outlet, as presented in Table 2. We utilize the Twitter API to collect tweets related to the COVID-19 pandemic from these accounts. The tweets are identified through contextual annotations provided by Twitter itself [41], a tool that employs artificial intelligence and manual annotations to determine the topic of posts on

**Table 2.** Statistics of the twitter accounts of the selected media outlets.

| Media | Account | Created at | Followers Count | Tweets Count | COVID Tweets Count | Replies Count |
|-------|---------|-----------|-----------------|--------------|--------------------|---------------|
| Globo News | g1 | 2007-09-11 | 13,503,670 | 728,050 | 13,636 | 1,139,807 |
| Folha de S. Paulo | folha | 2008-04-30 | 8,138,207 | 554,220 | 20,558 | 872,361 |
| UOL | UOLNoticias | 2008-04-30 | 4,208,743 | 526,127 | 25,628 | 681,007 |
| Estadão | Estadao | 2007-10-08 | 7,201,376 | 453,548 | 16,890 | 585,150 |
| CNN Brasil | CNNBrasil | 2016-12-14 | 1,551,680 | 89,129 | 24,262 | 575,504 |
| O Globo | JornalOGlobo | 2009-07-06 | 6,911,678 | 442,766 | 17,682 | 454,334 |
| Veja | VEJA | 2008-11-28 | 8,768,850 | 339,102 | 8,042 | 202,266 |
| BBC News | bbcbrasil | 2007-02-23 | 3,092,022 | 254,376 | 8,504 | 108,766 |
| Record News | portalR7 | 2009-08-13 | 4,950,294 | 500,451 | 12,035 | 89,059 |
| Jornal Extra | jornalextra | 2009-03-17 | 1,094,180 | 211,136 | 8,088 | 74,143 |
| BandNews | radiobandnewsfm | 2009-03-25 | 1,511,125 | 200,237 | 7,229 | 46,498 |
| Terra.com.br | Terra | 2009-10-08 | 3,011,485 | 199,083 | 4,620 | 20,414 |
| TV SBT | SBTonline | 2009-02-11 | 5,018,484 | 109,309 | 2,107 | 12,395 |

the social network. In total, 169,281 news-related tweets and 4,861,704 replies on these news items were collected, originated from 971,176 distinct users.

**Topic modeling** We employ a technique called Latent Dirichlet Allocation (LDA) [5] to identify the topics present in the news tweets. LDA is a statistical method that discovers latent topics from a corpus of documents. Before training, we exclude stopwords (i.e., very common words such as articles, prepositions, and others), web links, punctuation marks, numbers, and named entities of geographical type (e.g., city names, countries). Additionally, other terms that are either too common (i.e., appearing in more than one-third of the documents) or too rare (i.e., appearing in less than 0.05% of the documents) are removed. We also create tokens from bigrams — pairs of words that occur frequently together. At the end, the dictionary have 56,444 unique tokens, used in 169,281 documents, with an average of 7.56 tokens per document. We use the implementation of the MAchine Learning for LanguagE Toolkit (MALLET) library [20] for the LDA algorithm to topic modeling.

We identify the appropriate number of topics (as the LDA input) by analyzing the coherence scores ranging from 3 to 39, with a step of 3. The coherence of a topic, used as an indicator for topic quality, is based on the distributional hypothesis that words with similar meaning tend to co-occur within a similar context [39]. The coherence score of a model is the average of the coherence scores of all topics. A coherence score of 0.38 was achieved for 12 topics, in our dataset.

**Statistical Analysis** We use the chi-square test of independence [21] to assess whether a pair of categorical variables (e.g., emotion and media outlet) are related (alternative hypothesis) or independent (null hypothesis). This is a non-parametric test, appropriate for our dataset because some of the variables (i.e., emotion, topic, and media outlet) are nominal in nature. By analyzing the

test statistic and the degrees of freedom, we can determine the p-value. When the p-value is greater than the significance level ($p > 0.05$), we do not reject the null hypothesis, indicating insufficient evidence to conclude that the categorical variables are related. Additionally, the effect size measures the strength of the relationship between the variables; larger effect sizes suggest a stronger association.

We also use the Augmented Dickey-Fuller (ADF) Test [19] to assess whether a time series (e.g., the monthly prevalence of emotion in replies) is stationary (alternative hypothesis) or non-stationary (null hypothesis). A non-stationary time series exhibits some time-dependent structure (e.g., variance that changes over time). When the test statistic exceeds the critical value at a 5% significance level, this is reflected in the p-value ($p > 0.05$), leading us not to reject the null hypothesis, indicating insufficient evidence to conclude that the time series is stationary.

## 4  Results

In this work, our aim is to train a model for Emotion Detection (ED) in Brazilian Portuguese. We useCchatGPT for dataset translation, enabling the model to classify social media comments within the scope of Ekman's basic emotions model. We then utilize this model to analyze the distribution of emotions in replies to Brazilian news tweets related to the COVID-19 pandemic, considering variations related to the news outlet, the topic of the news, and the time of publication.

### 4.1  Emotion detection model

Table 3 presents the F1-score results for each emotion detected by the BERT model used in the GoEmotions study [9], as well as the performance achieved with our model for the GoEmotions-BR corpus. Although the results are not directly comparable, since they were calculated on different corpora, they provide an indicative of expected outcome. Our results demonstrate a similar order of magnitude and consistency across emotion classes, indicating that the developed model satisfactorily meets the conditions for analyzing emotion distribution in responses to COVID-19-related news tweets in Brazil.

### 4.2  Emotion distribution

We conducted a multi-stage statistical analysis to investigate how the distribution of emotions varies according to different factors. Firstly, we examine the global distribution of emotions in replies to news tweets, how reply volume varies over time, and how the distribution of emotions also changes over time. Subsequently, we used Latent Dirichlet Allocation (LDA) to identify latent topics within the corpus of COVID-19-related news tweets. Then, we assess how emotions are distributed among these topics. Finally, we evaluate how emotions are

**Table 3.** F1-score comparison of Ekman Emotion for GoEmotions and BERTimbau-Ekman-BR models.

| Emotion | GoEmotions | BERTimbau-Ekman-BR |
|---|---|---|
| anger | 0.57 | 0.50 |
| disgust | 0.53 | 0.49 |
| fear | 0.68 | 0.60 |
| joy | 0.82 | 0.80 |
| sadness | 0.59 | 0.54 |
| surprise | 0.61 | 0.55 |
| neutral | 0.66 | 0.64 |
| macro-average | 0.64 | 0.59 |
| std | 0.10 | 0.11 |

**Table 4.** Summary of chi-square test of independence results for multiple pairs of categorical variables. The test statistics is presented in thousands. A p-value lower than 5% rejects the null hypothesis that the two variables are independent. The table presents the rule of thumb over Cohen's $\omega$ measure [7] as the Effect Size result.

| Variables | Test Statistic | P-Value | Effect Size | Test Result |
|---|---|---|---|---|
| Emotion x Topic | 35 | 0.0% | Small | Reject the null hypothesis |
| Emotion x Media | 26 | 0.0% | Small | Reject the null hypothesis |

distributed among different news outlets. With that, we are able to map different scenarios impacting the distribution of emotions in replies, and statistically assess their impact. The results for multiple pairs of variables are summarized in Table 4, along with the effect size estimated using Cohen's $\omega$ [7]. Both the topic and the news outlet are hypothesized to influence the distribution of emotions, as the statistical test fails to reject the null hypothesis that they are independent in both cases. However, the effect in both instances is small.

**Distribution over time** In total, 4,861,704 replies to news tweets were collected over a three-year period, from January 2020 to December 2022. The volume of replies varied considerably throughout the observed period, peaking at over 300,000 posts per month in the first quarters of 2020 and 2021, but becoming virtually irrelevant from February 2022 onward. Conversely, Figure 1 shows the distribution of emotions across the entire corpus of news tweet replies. Just under half (40.9%) of the replies exhibit no detectable emotion, categorized as neutral. Among the replies with detectable emotions, anger (21.1%) emerges as the most prevalent, followed by joy (17.2%), surprise (11.6%), sadness (6.7%), disgust (1.3%), and fear (1.2%), which represents the least prevalent emotion.

Furthermore, Figure 2 presents the distribution of emotions over time in replies. This figure considers only the replies with detectable emotions, thus excluding those categorized as neutral. Over the course of time, slightly more
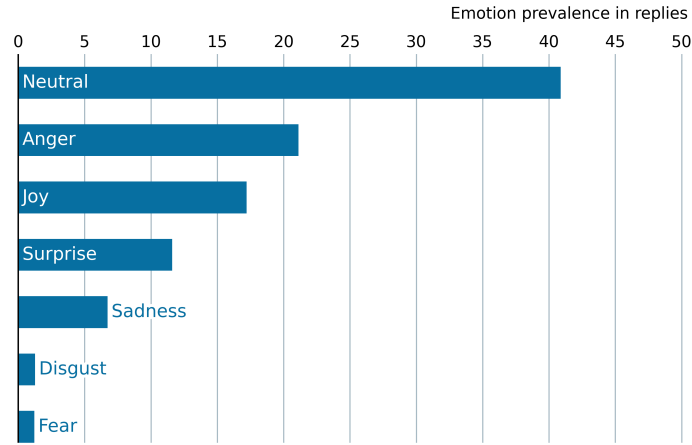
Emotion prevalence in replies



**Fig. 1.** The prevalence of emotions in user replies.
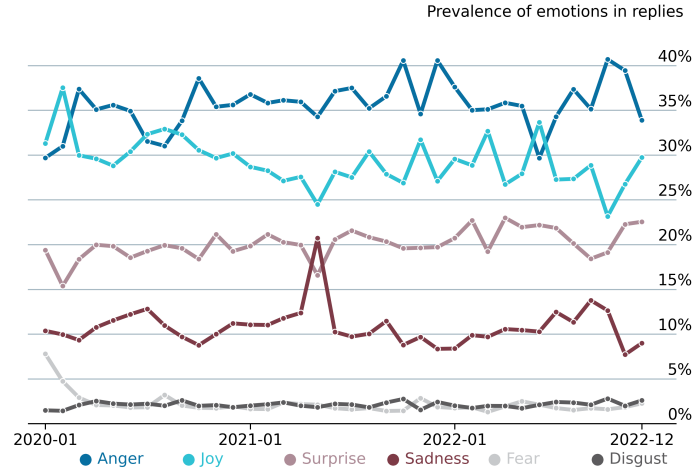
Prevalence of emotions in replies



**Fig. 2.** Monthly distribution of emotions in the corpus.

than a third of the replies with detectable emotions were associated with anger, while just under a third were linked to joy. The remainder was distributed among the other emotions, with sadness experiencing a significant peak in May 2021, surpassing 20% of the replies with detectable emotions.

Table 5 indicates that joy and neutral are likely the only emotions that are non-stationary, whereas anger, disgust, fear, surprise, and sadness are likely stationary. This suggests that the monthly prevalence of these emotions in user replies is not dependent on time. A more negative test statistic indicates a

**Table 5.** Augmented Dickey-Fuller test for monthly prevalence of emotions in user comments. A p-value higher than 5% fails to reject the null hypothesis that the time series of the daily amount of user comments with each emotion is non-stationary, thus indicating that the time series probably have time-dependent structures.

| Emotion | Test Statistics | P-Value | Test Result |
|---------|-----------------|---------|-------------|
| Anger | -4.7 | 0.0% | Reject the null hypothesis |
| Disgust | -5.8 | 0.0% | Reject the null hypothesis |
| Fear | -4.0 | 0.1% | Reject the null hypothesis |
| Joy | -1.7 | 41.1% | Fails to reject the null hypothesis |
| Sadness | -4.5 | 0.0% | Reject the null hypothesis |
| Surprise | -3.1 | 2.6% | Fails to reject the null hypothesis |
| Neutral | -3.6 | 8.9% | Reject the null hypothesis |

stronger rejection of the null hypothesis. Consequently, we can infer with greater confidence that joy is non-stationary, while disgust is stationary.

**Distribution over topics** Using LDA, we identify 12 topics in the news tweets, shown in Table 6. Each topic is represented by the ten most related words, a sample tweet, its prevalence among the news tweets, coherence score, and label (e.g., Cases & Deaths"). For instance, the first topic is more likely to be chosen by the LDA model if a news tweet contains keywords like coronavírus," saúde," or pacientes." This topic has a coherence score of 38.94%, appears in 15% of news tweets, and is labeled as "Impact on health system."

The chi-square test ($\chi^2$) shows a significant relationship between topics and the prevalence of emotions in user comments (Table 4), although with a small effect size. Table 7 provides $\chi^2$ results and the prevalence of emotions in comments according to the topic. A higher $\chi^2$ value indicates a greater deviation from the mean (e.g., anger in responses to vaccination news).

The highest $\chi^2$ is observed for anger in replies to Pandemic Investigation news (41.9%), higher than the mean anger prevalence (35.7%). This topic also has the lowest joy (25.4%) and sadness (9.2%). Replies to cases & deaths news show the lowest anger (29.4%) but the highest sadness (16.3%), surprise (21.5%), and fear (3.5%). News on preventive measures have the second-highest sadness (14.3%) and lowest surprise (17.6%). Joy is prevalent in replies to almost any topic, with the highest rates in vaccine distribution (34.6%) and international impact news (32.3%)."

**Distribution over media outlet** Media outlet and emotion are also two categorical variables that hold a significant relationship, according to the $\chi^2$ test result (Table 4), although with a small effect size. Table 8 displays the prevalence of each emotion in user comments as well as the $\chi2$ value for a given media outlet. The coverage of UOL has the highest prevalence of anger (40.0%) and the lowest of joy (26.1%) among user replies. Meanwhile, the highest prevalence of sadness (13.6%) and the lowest of surprise (18.0%) are observed in user replies to

**Table 6.** The topics identified by LDA.

| Keywords | Sample Tweets | Preva-lence | Coher-ence | Label |
|---|---|---|---|---|
| coronavírus, saúde, pacientes, dias, pessoas, doença, hospital, teste, médico, via_user | O enfermeiro Marcos (. . . ) contraiu doença e acabou internado na mesma UTI onde trabalha. | 0.15 | 38.94 | Impact on health system |
| bolsonaro, cpi, presidente, disse, pazuello, afirmou, user, stf, queiroga, governo_federal | Pazuello deve afirmar à PF em depoimento que recebeu pedido do presidente Bolsonaro de investigar supostas irregularidades (. . . ) | 0.13 | 53.04 | Pandemic investigation |
| vacinas, brasil, oms, mundo, afirma, países, testes, índia, esta-dos_unidos, ministério | Covid-19: 172 países já aderiram a acordo mundial de vacinas; OMS diz ter iniciado diálogo com autoridades russas | 0.10 | 34.77 | Vaccination debate |
| pandemia, brasil, ano, semana, será, alta, brasileiros, apesar, abril | Desigualdade bate recorde na pandemia e renda média cai a pior nível desde 2012 | 0.09 | 25.93 | Economic impact |
| quarentena, bolsonaro, anos, más-cara, fala, morre, gt, isolamento, casa, vídeo | Alexandre Garcia se descuida com máscara em evento com Bolsonaro. | 0.09 | 31.45 | Preventive measures |
| sp, governo, rio, lockdown, estados, paulo, doria, volta, rj, anuncia | Confira as novas regras da prefeitura para reabertura do Rio: do comércio e salões de rua a escolas e bares | 0.08 | 27.37 | Governmental actions |
| vacina, ministério_saúde, anvisa, coronavac, pfizer, crianças, fiocruz, doses, dose, butantan | O Instituto Butantan enviou (. . . ) uma so-licitação à Anvisa de início dos testes clíni-cos da vacina contra a Covid-19 Butanvac | 0.08 | 50.48 | Vaccine pro-duction |
| coronavírus, eua, país, china, causa, trump, onda, itália, surto, http_via | Coronavírus na Itália: como país conseguiu conter segunda onda de contágios e se tornou exceção na Europa | 0.06 | 34.71 | International impact |
| casos, mortes, brasil, dia, brasil_registra_mortes, país, horas, quase, mortos, doença | Brasil ultrapassa mil mortes e chega a quase 300 mil casos de Covid em 24 h. (. . . ) | 0.08 | 64.20 | Cases & Deaths |
| pandemia, combate, du-rante_pandemia, crise, devido, auxílio_emergencial, medidas, empresas | Projeto tem como objetivo criar um fundo emergencial para técnicos e artistas do seg-mento teatral afetados pela pandemia | 0.05 | 23.14 | Political re-sponse |
| vacinação, paulo, rio, governo, hoje, população, cidade, imuniza-ção, segundafeira | São Paulo inicia vacinação de pessoas de 72 a 74 anos na segunda-feira (22) | 0.04 | 36.20 | Vaccine dis-tribution |
| coronavírus, estudo, pessoas, reino_unido, ômicron, caso, vírus, variante, risco, manaus | Pesquisadores da UFMG divulgaram de-scoberta de 18 mutações no Sars-Cov-2 (. . . ) | 0.04 | 35.69 | Research & variants |

**Table 7.** Prevalence and chi-square values (in thousands) of emotions in user comments per topic. A higher chi-square value indicates a more significant deviation from the expected (random) prevalence. For example, anger appears in 20.5% of comments on cases & deaths news, with a chi-square value of $5.9x10^3$.

| Topic | Anger | Disgust | Fear | Joy | Sadness | Surprise |
|---|---|---|---|---|---|---|
| Impact on health system | 33.5% (0.6) | 2.0% (0.0) | 2.2% (0.1) | 30.2% (0.2) | 11.5% (0.0) | 20.5% (0.2) |
| Pandemic investigation | **41.9% (6.2)** | **2.5% (0.5)** | 1.9% (0.1) | 25.4% (3.1) | 9.2% (2.6) | 19.2% (0.1) |
| Vaccination debate | 35.4% (0.0) | 1.8% (0.1) | 2.0% (0.0) | 29.7% (0.0) | 10.3% (0.3) | 20.8% (0.1) |
| Economic impact | 34.5% (0.2) | 2.1% (0.0) | 2.3% (0.1) | 29.9% (0.0) | 11.4% (0.0) | 19.8% (0.0) |
| Preventive measures | 34.9% (0.0) | **2.5% (0.5)** | 1.8% (0.1) | 28.9% (0.1) | **14.3% (4.5)** | **17.6% (0.4)** |
| Governmental actions | 35.5% (0.0) | 2.0% (0.0) | 1.6% (0.2) | 30.4% (0.1) | 9.4% (0.7) | 21.2% (0.2) |
| Vaccine production | 32.4% (0.7) | 1.6% (0.3) | 2.0% (0.0) | **32.3% (0.8)** | 13.6% (1.0) | 18.2% (0.2) |
| International impact | 32.0% (0.9) | 1.8% (0.1) | 2.3% (0.1) | 32.3% (0.2) | 10.1% (0.3) | 21.5% (0.1) |
| Cases & Deaths | 29.4% (1.4) | 2.3% (0.0) | **3.5% (1.4)** | 27.0% (0.2) | 16.3% (2.9) | 21.5% (0.3) |
| Political response | 37.7% (0.1) | 2.2% (0.0) | 1.6% (0.2) | 29.5% (0.0) | 9.8% (0.3) | 19.2% (0.0) |
| Vaccine distribution | 32.6% (0.3) | 1.4% (0.2) | 1.7% (0.1) | 34.6% (0.7) | 9.3% (0.3) | 20.5% (0.0) |
| Research & variants | 36.3% (0.0) | 1.8% (0.1) | 2.6% (0.1) | 29.3% (0.0) | 11.1% (0.0) | 18.8% (0.0) |
| **Mean** | **35.7%** | **2.1%** | **2.0%** | **29.1%** | **11.4%** | **19.6%** |

the news from g1, with a significant deviation from the global mean. The highest prevalence of joy is observed in user replies to BBC News Brasil (32.9%), and surprise reach the highest prevalence in user replies to the news published by SBT (22.3%).

**Table 8.** Prevalence and chi-square values (in thousands) of emotions in user comments per media outlet. A higher chi-square value indicates a more significant deviation from the expected prevalence. For example, anger appears in 29.7% of comments on BBC News Brasil, with a chi-square value of $1.1x10^3$.

| Media outlet | Anger | Disgust | Fear | Joy | Sadness | Surprise |
|---|---|---|---|---|---|---|
| BBC News Brasil | 29.7% (1.1) | 2.0% (0.0) | 3.0% (0.2) | 32.9% (0.1) | 12.7% (0.0) | 19.7% (0.0) |
| CNN Brasil | 33.4% (0.2) | 1.8% (0.1) | 1.8% (0.1) | 28.0% (0.0) | 13.1% (1.4) | **21.9% (1.5)** |
| Estadão | 37.0% (0.1) | 2.1% (0.0) | 1.8% (0.1) | 28.5% (0.1) | 10.1% (0.5) | 20.4% (0.1) |
| Folha de S.Paulo | 39.0% (1.5) | **2.5% (0.3)** | 1.8% (0.1) | 27.7% (0.4) | 10.0% (0.9) | 19.0% (0.1) |
| Jornal Extra | 30.0% (0.5) | 1.6% (0.1) | 1.9% (0.0) | 36.9% (0.7) | 11.2% (0.0) | 18.4% (0.1) |
| Jornal O Globo | 35.8% (0.0) | 1.9% (0.1) | 2.2% (0.0) | 29.8% (0.0) | 9.8% (0.7) | 20.5% (0.0) |
| Portal R7.com | 33.6% (0.1) | 1.8% (0.0) | 2.0% (0.0) | 30.9% (0.0) | 9.3% (0.2) | 22.3% (0.1) |
| Rádio BandNews FM | 31.7% (0.3) | 1.8% (0.0) | 2.5% (0.0) | 30.2% (0.0) | 12.3% (0.0) | 21.6% (0.0) |
| SBT | 36.0% (0.0) | 3.0% (0.0) | 1.6% (0.0) | 30.6% (0.0) | 12.0% (0.0) | 16.8% (0.0) |
| Terra | 36.4% (0.0) | 2.1% (0.0) | 1.9% (0.0) | 27.7% (0.0) | 11.0% (0.0) | 20.9% (0.0) |
| UOL Notícias | **40.0% (1.7)** | 2.3% (0.1) | 1.9% (0.1) | 26.1% (1.5) | 10.2% (0.5) | 19.5% (0.0) |
| VEJA | 36.2% (0.0) | 2.1% (0.0) | 1.9% (0.0) | 29.7% (0.0) | 10.2% (0.2) | 19.9% (0.0) |
| g1 | 32.4% (1.4) | 2.1% (0.0) | **2.4% (0.6)** | **31.5% (1.9)** | **13.6% (3.6)** | 18.0% (0.6) |
| **Mean** | **35.7%** | **2.1%** | **2.0%** | **29.1%** | **11.4%** | **19.6%** |

# 5   Conclusions and future work

In this study, we developed a specialized Emotion Detection (ED) model for Brazilian Portuguese social media comments. While GoEmotions [9] is widely used, there have been no prior efforts to create emotion classifiers in Brazilian Portuguese that preserve the original social media writing style. To address this, we translated GoEmotions into GoEmotions-BR, with chatGPT-3.5-turbo effectively capturing nuanced expressions. The BERTimbau-base model, fine-tuned on GoEmotions-BR, enabled the detection of Ekman's six basic emotions and neutral emotion in Portuguese.

Our model achieved a F1-score in the 0.5-0.6 range, close to state-of-the-art results, with consistent performance across emotions, indicating evenly distributed errors and minimal bias. Although there is room for improvement, the model's performance is significantly better than chance, given the number of emotion classes.

We analyzed factors influencing emotion distribution in replies to COVID-19-related news, using 169,281 tweets and 4,861,704 replies. LDA identified the top 12 topics, and chi-squared tests confirmed that emotion distribution depended on both the news topic and the publishing outlet. The ADF test revealed non-stationarity in joy prevalence, while other emotions were stationary.

Most replies were neutral, with anger (21.1%) being the most prevalent identifiable emotion, followed by joy (17.2%) and surprise (11.6%). The 2021 Brazilian pandemic inquiry had the highest prevalence of anger (41.9%), preventive measures news led to the most sadness (14.3%), and vaccine production news elicited the most joy (32.3%). UOL Notícias had the highest anger (40.0%) and lowest joy (26.1%), while g1 had the highest sadness (13.6%).

Future research should focus on creating ED datasets that reflect Brazilian cultural nuances and evaluating emotion classifiers with other language models, such as Albertina [33], peLLE [24], and Sabiá [31]. Additionally, comparing the results on GoEmotions-BR with other strategies, like bag-of-words, could help assess the proposed model's quality. Finally, the methodology could be extended to other tasks, such as hate speech detection.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

**Data and Models Availability** The data and models developed in this study will be made available upon request.

# References

1. Akhtar, M.S., Chauhan, D.S., Ghosal, D., Poria, S., Ekbal, A., Bhattacharyya, P.: Multi-task learning for multi-modal emotion recognition and sentiment analysis. arXiv preprint arXiv:1905.05812 (2019)
2. Al Maruf, A., Khanam, F., Haque, M.M., Jiyad, Z.M., Mridha, F., Aung, Z.: Challenges and opportunities of text-based emotion detection: A survey. IEEE Access (2024)
3. Amin, M.M., Mao, R., Cambria, E., Schuller, B.W.: A wide evaluation of chatgpt on affective computing tasks. arXiv preprint arXiv:2308.13911 (2023)
4. Bernardi, F.A., Lima, V.C., Rijo, R.P.C.L., Alves, D.: Mais do que palavras: uma análise das emoções brasileiras durante a covid-19. Revista Ibérica de Sistemas e Tecnologias de Informação (E40), 526–541 (2021)
5. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. the Journal of machine Learning research **3**, 993–1022 (2003)
6. Cheong, M., Lee, V.C.: A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via twitter. Information Systems Frontiers **13**, 45–59 (2011). https://doi.org/10.1007/s10796-010-9273-x
7. Cohen, J.: Statistical power analysis for the behavioral sciences. Routledge (2013)
8. Cui, Y., Yang, Z., Yao, X.: Efficient and effective text encoding for chinese llama and alpaca. arXiv preprint arXiv:2304.08177 (2023)
9. Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., Ravi, S.: Goemotions: A dataset of fine-grained emotions. arXiv preprint arXiv:2005.00547 (2020)

10. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
11. Ekman, P.: An argument for basic emotions. Cognition & emotion **6**(3-4), 169–200 (1992)
12. Fellbaum, C.: Wordnet. In: Theory and applications of ontology: computer applications, pp. 231–243. Springer (2010)
13. Gill, A.J., French, R.M., Gergle, D., Oberlander, J.: Identifying emotional characteristics from short blog texts. In: Proceedings for the 30th Annual Meeting of the Cognitive Science Society. pp. 2237–2242. Cognitive Science Society (2008)
14. Hammes, L.O.A., de Freitas, L.A.: Utilizando bertimbau para a classificação de emoções em português. In: Anais do XIII Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana. pp. 56–63. SBC (2021)
15. Hung, L.P., Alias, S.: Beyond sentiment analysis: A review of recent trends in text based sentiment analysis and emotion detection. Journal of Advanced Computational Intelligence and Intelligent Informatics **27**(1), 84–95 (2023)
16. Kusal, S., Patil, S., Choudrie, J., Kotecha, K., Vora, D., Pappas, I.: A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection. Artificial Intelligence Review **56**(12), 15129–15215 (2023)
17. Larcher, C., Piau, M., Finardi, P., Gengo, P., Esposito, P., Caridá, V.: Cabrita: closing the gap for foreign languages. arXiv preprint arXiv:2308.11878 (2023)
18. Lwin, M.O., Lu, J., Sheldenkar, A., Schulz, P.J., Shin, W., Gupta, R., Yang, Y.: Global sentiments surrounding the covid-19 pandemic on twitter: analysis of twitter trends. JMIR public health and surveillance **6**(2), e19447 (2020)
19. MacKinnon, J.G.: Approximate asymptotic distribution functions for unit-root and cointegration tests. Journal of Business & Economic Statistics **12**(2), 167–176 (1994)
20. McCallum, A.K.: Mallet: A machine learning for language toolkit (2002), http://mallet.cs.umass.edu
21. McHugh, M.L.: The chi-square test of independence. Biochemia medica **23**(2), 143–149 (2013)
22. Mehrabian, A.: Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies (1980)
23. Mehrabian, A.: Communication without words. In: Communication theory, pp. 193–200. Routledge (2017)
24. de Mello, G.L., Finger, M., Carpi, M.d.M., Jose, M.M., Domingues, P.H., Cavalim, P., et al.: Pelle: Encoder-based language models for brazilian portuguese based on open data. arXiv preprint arXiv:2402.19204 (2024)
25. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 (2013)
26. Newman, N., Fletcher, R., Eddy, K., Robertson, C.T., Nielsen, R.K.: Digital news report 2023 (2023)
27. Oliveira, F.B., Haque, A., Mougouei, D., Evans, S., Sichman, J.S., Singh, M.P.: Investigating the emotional response to covid-19 news on twitter: a topic modeling and emotion classification approach. IEEE Access **10**, 16883–16897 (2022)
28. O'Connor, B., Balasubramanyan, R., Routledge, B., Smith, N.: From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. Proceedings of the International AAAI Conference on Web and Social Media (ICWSM) **4**(1), 122–129 (May 2010). https://doi.org/10.1609/icwsm.v4i1.14031

29. Pasqualotti, P.R., Vieira, R.: Wordnetaffectbr: uma base lexical de palavras de emoções para a língua portuguesa. Revista Novas Tecnologias na Educação **6**(1) (2008)
30. Picard, R.W.: Affective computing. MIT press (2000)
31. Pires, R., Abonizio, H., Almeida, T.S., Nogueira, R.: Sabiá: Portuguese large language models. In: Brazilian Conference on Intelligent Systems. pp. 226–240. Springer (2023)
32. Plutchik, R.: Emotions: A general psychoevolutionary theory. Approaches to emotion **1984**, 197–219 (1984)
33. Rodrigues, J., Gomes, L., Silva, J., Branco, A., Santos, R., Cardoso, H.L., Osório, T.: Advancing neural encoding of portuguese with transformer albertina pt. In: EPIA Conference on Artificial Intelligence. pp. 441–453. Springer (2023)
34. Russell, J.A.: A circumplex model of affect. Journal of personality and social psychology **39**(6), 1161 (1980)
35. Scherer, K.: Profiles of emotion-antecedent appraisal: Testing theoretical predictions across cultures. Cognition & Emotion **11**(2), 113–150 (1997)
36. Silva, L.G., Caseli, H.M.: Aspect-based sentiment analysis in comments on political debates in portuguese: evaluating the potential of chatgpt. PROPOR 2024 p. 312 (2024)
37. Souza, F., Nogueira, R., Lotufo, R.: Bertimbau: pretrained bert models for brazilian portuguese. In: Intelligent Systems: 9th Brazilian Conference, BRACIS 2020, Rio Grande, Brazil, October 20–23, 2020, Proceedings, Part I 9. pp. 403–417. Springer (2020)
38. Strapparava, C., Mihalcea, R.: Learning to identify emotions in text. In: Proceedings of the 2008 ACM symposium on Applied computing. pp. 1556–1560 (2008)
39. Syed, S., Spruit, M.: Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). pp. 165–174 (2017). https://doi.org/10.1109/DSAA.2017.61
40. Sykora, M., Jackson, T., O'Brien, A., Elayan, S.: Emotive ontology: Extracting fine-grained emotions from terse, informal messages (2013)
41. Twitter, I.: Twitter context annotations: List of entities. https://github.com/twitterdev/twitter-context-annotations (2022)
42. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
43. Wagner Filho, J.A., Wilkens, R., Idiart, M., Villavicencio, A.: The brwac corpus: A new open resource for brazilian portuguese. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018) (2018)
44. Wake, N., Kanehira, A., Sasabuchi, K., Takamatsu, J., Ikeuchi, K.: Bias in emotion recognition with chatgpt. arXiv preprint arXiv:2310.11753 (2023)
45. Wang, X., Zheng, Q.: Text emotion classification research based on improved latent semantic analysis algorithm. In: Conference of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013). pp. 210–213. Atlantis Press (2013)
46. Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., Gao, S., Sun, Y., Ge, W., Zhang, W., et al.: A systematic review on affective computing: Emotion models, databases, and recent advances. Information Fusion **83**, 19–52 (2022)
47. Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., Zhu, T.: Twitter discussions and emotions about the covid-19 pandemic: Machine learning approach. Journal of medical Internet research **22**(11), e20550 (2020)