

Analisando as emoções dos tweets relacionadas à Covid-19 no Rio de Janeiro*

Gustavo F. L. Gonçalves¹, Antonio A. de A. Rocha¹, Aline Paes¹

¹ Instituto de Computação (IC)
Universidade Federal Fluminense (UFF) – Niteroi, RJ – Brazil

gustavoferreira@id.uff.br; {arocha, alinepaes}@ic.uff.br

Abstract. *This paper aims to analyze tweets posts related to Covid-19 to show which were the predominant users emotions. To that we collect tweets related to Rio de Janeiro and yield statistical analyzes and emotion classifiers. In addition, the paper also provide insights of the subjects discussed by the users. Furthermore, statistical analyses help on showing how the emotions change according to specific events. The tools developed here helps to understand the behaviors and emotions of Twitter users, resulting in evidence that can be useful in similar catastrophic situations.*

Resumo. *Este trabalho tem como objetivo analisar postagens de tweets relacionados à Covid-19 para mostrar quais foram as emoções predominantes dos usuários. Para isso coletamos tweets relacionados ao Rio de Janeiro e produzimos análises estatísticas e classificadores de emoções. Além disso, o artigo também traz insights sobre os assuntos discutidos pelos usuários e como as emoções mudaram de acordo com eventos específicos. As ferramentas aqui desenvolvidas ajudam a compreender os comportamentos e emoções dos usuários do Twitter, resultando em evidências que podem ser úteis em situações catastróficas semelhantes.*

1. Introdução

Desde o fim do ano de 2019, o mundo inteiro se deparou com um problema grave: a Covid-19, que por ser uma doença altamente transmissível, rapidamente se espalhou pelo mundo em 2020, chegando ao status de pandemia em 11 de Março de 2020, atingindo primeiramente a Europa e, após isso, os outros continentes do planeta. Diversas ações foram tomadas para impedir o avanço da doença, tais como: o uso de máscaras e álcool em gel, implementação de isolamento social e *lockdowns*, na tentativa de frear a quantidade de casos. Apesar de não haver um consenso geral entre todos os países, existiu e ainda existe um esforço global para que todas as pessoas possam voltar à vida normal o mais rápido possível.

Com o complexo cenário vivido por toda população mundial, entender como as pessoas estão se relacionando, seus estados emocionais e como a pandemia está afetando cada uma das pessoas é uma tarefa importante e vital para que possamos ajudar uns aos outros e nos prepararmos para eventos futuros que possam ser similares ao que estamos vivendo, sendo esse o objetivo principal desse artigo. Embora existam diversos trabalhos relacionados à área de Aprendizado de Máquina que aplicam

*Esse trabalho foi parcialmente apoiado pela FAPESP (processo 2015/24144-7).

soluções como análise de emoções ou sentimentos [Li et al. 2020, Nemes and Kiss 2021, Brum et al. 2020, Rustam et al. 2021, Müller et al. 2020], ainda poucos trabalhos aplicam tais técnicas no contexto da pandemia da Covid-19 no Brasil, principalmente por ser uma situação muito recente para todos nós. Com isso em mente, o trabalho promovido nesse estudo busca aplicar técnicas de Aprendizado de Máquina para dados obtidos em redes sociais durante a pandemia da Covid-19, de forma a analisar o comportamento dos usuários da amostra de dados coletada.

A análise de emoções desenvolvida considera dados coletados do Twitter na região do estado do Rio de Janeiro. Os dados foram transformados por diversas formas para que servissem como entrada para diferentes classificadores. Os classificadores foram comparados para analisar a melhor opção para esse problema. Além disso, foram gerados múltiplos indicadores que exibem o comportamento dos usuários ao longo do tempo no Twitter durante os primeiros meses da pandemia da Covid-19.

O restante deste artigo está organizado da seguinte forma: a Seção 2 possui os trabalhos relacionados ao estudo, a Seção 3 descreve as etapas desenvolvidas para a análise de emoções, a Seção 4 apresenta os resultados obtidos e a Seção 5 conclui o estudo.

2. Trabalhos Relacionados

Diversos estudos têm se focado na construção de classificadores para identificar sentimentos de usuários do Twitter. Nessa Seção, incluímos artigos que focam na tarefa de análise de sentimentos para dados relacionados à pandemia da Covid-19.

O trabalho mais similar ao apresentado neste artigo foi realizado por Li et al. [Li et al. 2020]. O artigo analisou sentimentos para uma base de dados com mais de oito milhões de tweets originados de diversos países do mundo, como os Estados Unidos, Índia, Inglaterra e Brasil, com datas de publicação entre os dias 24 e 26 de Março de 2020. Os autores treinaram um classificador a partir do modelo de linguagem BERT [Devlin et al. 2019] para a análise de emoções. Cada tweet poderia ser classificado em ao menos uma das seguintes emoções: raiva, ansiedade, desgosto, medo, alegria, tristeza, surpresa e confiança. A tarefa foi modelada como classificação de rótulo único e classificação multirrótulo.

Brum et al. [Brum et al. 2020] apresentaram uma análise de tweets em português que estavam relacionados com a pandemia da Covid-19, comparando os comportamentos dos usuários ao longo do tempo. Os tweets foram coletados durante o período de 23 de abril e 02 de julho de 2020. O trabalho incluiu várias análises, considerando o volume de tweets ao longo do tempo, nuvens de palavras para tweets e retweets, para identificar temas relacionados ao uso de algumas palavras, distribuição geográfica dos tweets, uso de *URLs*, *hashtags* e menções a usuários, e os perfis dos usuários, para inferir gêneros para cada um dos usuários, encontrar usuários jornalísticos e o uso de *bots*¹.

Em [Rustam et al. 2021] comparou-se o desempenho de diversos classificadores de polaridade de tweets em inglês, onde cada um poderia ser classificado como positivo, neutro ou negativo. Os tweets são oriundos de um *dataset* disponibilizado no site IEEE DataPort, onde existe uma pontuação de emoção vinculada a cada tweet, calculada com a

¹Os *bots* (ou *social bots*) são perfis de usuários automatizados que atuam em mídias sociais com a finalidade de produzir/disseminar algum tipo de conteúdo ou realizar algum tipo de ação

biblioteca TextBlob [Loria 2018]. Para representar numericamente os tweets, foi usada a técnica de Bag-of-words com TF-IDF ou contagem e uma combinação entre essas duas. Os classificadores utilizados foram Random Forest, XGBoost, SVC, Extra Trees e Decision Tree, utilizando as medidas de acurácia, precisão, sensibilidade e F-1 para comparar o desempenho de cada classificador. As melhores acurácias obtidas foram com o classificador Extra Trees.

Müller et al. [Müller et al. 2020] criaram o modelo CT-BERT, utilizando o modelo pré-treinado BERT-LARGE. Foram criados cinco conjuntos de treinamento multiclasse independentes entre si: Categoria de Covid-19, Sentimento sobre vacina, Posição relacionada ao uso de vacinas, Sentimentos no Twitter e Banco de Sentimentos de Stanford. O modelo CT-BERT foi treinado com tweets relacionados ao coronavírus obtidos a partir da plataforma Crowdbreaks. Pôde-se observar que os melhores resultados estavam relacionados aos datasets que tinham temáticas parecidas com os dados usados no treinamento (Categoria de Covid-19, Sentimento sobre vacina e Posição relacionada ao uso de vacinas). Mesmo assim, em todos os casos, houve uma melhora nas métricas obtidas com o CT-BERT quando comparados com as classificações feitas usando o modelo BERT-LARGE.

Os trabalhos citados possuem diversos pontos em comum com o presente artigo: a coleta de tweets através do uso de APIs, o uso de técnicas de transformação de dados como TF-IDF, Bag-of-Words e word embedding, a construção de um dataset com esses tweets, a construção de um classificador de usando diversos métodos (SVM, Logistic Regression, Random Forest, BERT, etc) e a comparação entre os desempenhos de cada classificador. O principal diferencial desse artigo é a classificação de emoções, ao invés de polaridade, para um conjunto de dados escritos em português, tendo como inspiração o trabalho desenvolvido em [Li et al. 2020]. Ademais, o classificador de emoções é aplicado em dados do estado do Rio de Janeiro, com análises conduzidas para compreender os dados coletados.

3. Metodologia

A metodologia proposta neste artigo envolve a utilização de técnicas distintas de Aprendizado de Máquina. Os tweets coletados foram restritos ao estado do Rio de Janeiro, usando a informação de localidade fornecida pelo usuário em seu perfil. Após, os tweets passaram por um processo de anotação manual de emoções, onde cada um dos tweets foi vinculado a uma emoção do estudo. Então, duas emoções foram escolhidas para criar um dataset menor, que passará pelo processo de treinamento. Os dados são pré-processados de forma a servirem de entrada para os métodos de aprendizado de máquina, usando a técnica clássica de Bag-of-words [Harris 1954, Scott and Matwin 1999], e usando word embeddings [Mikolov et al. 2013a], que são vetores densos de baixa dimensionalidade pré-treinados. Espera-se que os embeddings capturem a semântica das palavras, de forma que palavras com semântica similar fiquem próximas no espaço vetorial. Dois modelos são usados para capturar os embeddings dos tweets, um método que gera uma representação estática independente para cada palavra e um método contextualizado, que considera as demais palavras do tweet para obter a representação vetorial. Os classificadores são aprendidos usando um processo de validação cruzada, com as medidas preditivas sendo a média dos conjuntos de teste. Cada um desses processos será abordado em uma subseção própria no decorrer desse artigo.

3.1. Coleta de tweets

Para efetuar a coleta dos tweets, utilizou-se a API GetOldTweets². Um total de mais de 200.000 tweets foram coletados entre os dias 1 de março de 2020 e 23 de maio de 2020, inclusive. Esse período envolve os primeiros três meses da pandemia da Covid-19 no Brasil. A busca dos tweets foi feita a partir de palavras-chave como “**coronavirus**”, “**covid-19**”, “**isolamento**”, “**pandemia**”, e “**quarentena**” que são termos diretamente relacionados à situação vivida pelo mundo inteiro com a Covid. Além disso, as *tags* de localização definida nos perfis de cada usuário do Twitter foram utilizadas para filtrar tweets que não pertencessem ao estado do Rio de Janeiro.

3.2. Rotulação dos tweets

As classes escolhidas para a rotulação dos tweets são as mesmas emoções utilizadas em [Li et al. 2020], ou seja, Raiva, Ansiedade, Desgosto, Medo, Alegria, Tristeza, Surpresa, Confiança, com a adição de duas opções (*Outra* e *Neutra*), para o caso de um tweet que não possa ser relacionado com nenhuma das demais emoções. As oito emoções foram selecionadas de acordo com a teoria da Roda das Emoções, criada por Plutchik [Plutchik 1984]. Essa é uma teoria da psicologia que define as emoções básicas do ser humano, cunhando as emoções derivadas às básicas e as emoções opostas entre si. O estudo promovido aqui neste artigo possui algumas diferenças com a Roda das Emoções, uma vez que Nojo e Antecipação não foram consideradas em nosso estudo e aparecem na Roda de Emoções. No entanto, essas emoções podem ser correlacionadas com Desgosto e Ansiedade, que foram consideradas aqui.

Para realizar a anotação manual de tweets, foi criado um sistema responsável por buscar tweets aleatórios do conjunto de dados coletado e divulgá-los em um formulário. O formulário exibe 20 tweets por vez, permitindo ao usuário continuar a anotação, ou não. A divulgação desse formulário se deu em duas etapas distintas: expansão do dataset de tweets rotulados e refinamento dos tweets rotulados. Além disso, não foi estabelecido nenhum tipo de perfil de participantes da pesquisa, devido ao limitado número de participantes nesse processo de rotulação.

Na etapa de expansão, o objetivo era obter o número máximo possível de tweets rotulados pelos participantes, removendo tweets já rotulados da busca aleatória. Essa primeira etapa gerou cerca de 1300 tweets com apenas uma rotulação vinculada e com a participação de, aproximadamente, 65 participantes. A partir dessa anotação, os tweets com anotação “*Outra*” ou “*Neutra*” foram descartados.

Na etapa de refinamento ocorre um novo processo de anotação. Entretanto, apenas os tweets rotulados na etapa anterior são considerados, em uma tentativa de aumentar a quantidade de pessoas que anotam um mesmo tweet. Dessa maneira, as avaliações das emoções em cada tweet não ficam restritas a um único usuário, evitando uma análise com alto grau de subjetividade e aumentando o grau de confiança. Foram obtidas 1600 rotulações que, somadas às 1300 anteriores, resultou em 2900 rotulações no total. Para identificar a principal emoção de um tweet, considerou-se a emoção que foi mais selecionada pelos participantes da pesquisa. Em caso de empate, o tweet foi replicado para cada emoção que tivesse o mesmo número de anotação. Esses casos de empate ocorreram para

²Disponibilizado em: <https://github.com/Jefferson-Henrique/GetOldTweets-python>

as seguintes combinações, que são emoções exibidas em conjunto em diversas situações: (1) Raiva, Desgosto e Tristeza; (2) Desgosto, Medo e Tristeza; (3) Ansiedade e Medo; (4) Alegria e Surpresa; (5) Tristeza e Surpresa. Após esses processos, foram obtidos 700 tweets anotados com alguma emoção, descartando todos os tweets que obtiveram as emoções “Outra” ou “Neutra”.

3.3. Seleção de emoções

O problema foi modelado como multiclasse, considerando todas as emoções como possível rótulo do tweet. Em adição, o problema também foi modelado como um classificador binário, para facilitar a discriminação entre diferentes emoções. Assim, a partir do conjunto de tweets obtido em cada uma das etapas de rotulação, duas emoções foram selecionadas para modelar o problema como classificação binária, combinando as emoções duas a duas. Por exemplo, na combinação de Alegria e Tristeza, o dataset a ser usado envolve apenas os tweets de Alegria e Tristeza.

3.4. Pré-processamento e separação dos dados

Nessa etapa, ocorre o tratamento do conjunto de dados, excluindo elementos de um tweet que podem atrapalhar a classificação de emoções. Para tanto, foi utilizada as bibliotecas Ekphrasis [Baziotis et al. 2017] e paCy [Honnibal et al. 2020] para tratar os seguintes elementos dos tweets:

1. Remoção de links, nomes de usuários, e-mails, símbolo de retweet (*RT*) e símbolos de hashtags (#).
2. Conversão para letras minúsculas.
3. Remoção de valores numéricos isolados. Expressões compostas por letras e números não se encaixam nessa categoria.
4. Remoção de *Stop-words*, que são palavras que normalmente não agregam valor semântico.
5. Remoção das palavras-chave que foram utilizadas na etapa de coleta de tweets.

Técnicas como estemização e lematização não foram consideradas, uma vez que os embeddings pre-treinados contextualizados não requerem esse tipo de pré-processamento. Testes preliminares indicaram que os classificadores gerados com as demais representações também não se beneficiavam dessas técnicas.

Para treinar os classificadores, foi utilizado o procedimento de validação cruzada com cinco folds. Dessa forma, a cada iteração, o conjunto de treinamento é composto por 80% dos tweets e o conjunto de teste pelos 20% restantes.

3.5. Transformação da representação textual para uma representação numérica

Para que os tweets possam ser usados como entrada em métodos de aprendizado de máquina, é necessário representá-los em um formato numérico. Para tanto, foram usadas as seguintes técnicas de extração de características:

1. No método de Bag-of-Words com contagem um texto é representado como se fosse um “saco” de palavras, onde a ocorrência de cada palavra é usada como característica [Harris 1954] e a ordem das palavras não é levada em consideração. Para computar as contagens das palavras em um tweet foi utilizada a implementação fornecida na biblioteca sklearn [Pedregosa et al. 2011].

2. TF-IDF é uma medida estatística que indica a importância de uma palavra em um documento em relação a uma coleção de documentos [Roelleke and Wang 2008, Ko 2012]. Para aplicar essa técnica, utilizou-se a implementação disponibilizada na biblioteca sklearn [Pedregosa et al. 2011].
3. Word Embedding estático é uma representação de palavras no formato de vetores densos com valores reais, onde, intuitivamente, cada valor representa uma característica da palavra [Mikolov et al. 2013b]. Para a transformação dos tweets em embeddings, foi utilizado o conjunto de vetores de palavras em português fornecido pela biblioteca fastText [Joulin et al. 2017].
4. Word Embedding contextualizado. BERT é uma arquitetura de redes neurais para modelagem de linguagens baseada no componente *encoder* de um Transformer [Vaswani et al. 2017]. A partir do uso de mecanismos de auto-atenção, espera-se capturar as inter-relações entre as diferentes partes do texto que estão sendo processadas. Para tanto, o BERT utiliza duas tarefas auto-supervisionadas, a previsão de palavras que são mascaradas aleatoriamente (*Masked Language Model*) e a previsão se uma sentença é sequência lógica da sentença anterior (*Next Sequence Prediction*). Essas sentenças são concatenadas em estruturas maiores como frases e parágrafos para que a técnica funcione apropriadamente. Assim, o BERT é forçado a usar as palavras na proximidade da palavra mascarada, e o contexto das duas sentenças apresentadas ao modelo em sequência, para fazer a predição correspondente, aprendendo a inferir uma palavra a partir do contexto da sentença onde ela está inserida. Para esse estudo, foi utilizada a versão em português do modelo de linguagem do BERT, o BERTimbau [Souza et al. 2020].

3.6. Treinamento dos classificadores

Os classificadores usados nesse estudo são: Random Forest, kNN, Passive Aggressive, Naive Bayes, Gradient Boosting, XGB, MLP, Logistic Regression e SVC. Para encontrar os melhores hiperparâmetros para cada classificador, utilizou-se a busca *GridSearchCV* implementada na biblioteca SciKit-Learn. Essa é uma busca exaustiva feita a partir dos hiperparâmetros e de seus valores especificados como entrada. Os classificadores que não receberam esse tratamento foram o Naive Bayes, que teve seus parâmetros definidos como os valores padrão, e a camada de classificação anexada ao BERT, que teve seus parâmetros fixados a partir do estudo feito em [Li et al. 2020] e seu processo de fine-tuning realizado com conjuntos de treinamento correspondentes a cada dataset usado.

3.7. Avaliação dos classificadores

A principal métrica utilizada para avaliar as performances dos classificadores foi a acurácia. A acurácia representa a fração das predições feitas corretamente, podendo alcançar um valor máximo de 1, quando todos os dados são previstos corretamente, ou um valor mínimo de 0, quando nenhum dado é previsto corretamente. A acurácia é a soma dos exemplos classificados corretamente, dividida pela soma de todos os exemplos.

Em um problema de classificação binária, TP (*True Positive*) são os exemplos de uma classe *A* classificados pelo método como pertencente àquela classe, enquanto TN (*True Negative*) são os exemplos da classe *B* também classificados corretamente. FP (*False Positive*) são os exemplos da classe *B* que foram classificados como sendo da classe *A*, enquanto FN (*False Negative*) são os exemplos da classe *A* que foram classificados como sendo da classe *B*.

4. Resultados obtidos

Esta seção apresenta os resultados obtidos das análises desenvolvidas neste trabalho.

4.1. Análise de emoções

A Tabela 1 exibe os resultados obtidos para o problema de multi-classe que tenta distinguir as oito emoções. Esse dataset foi obtido na etapa de expansão dos dados da Seção 3.2. É possível notar que existem casos de *overfitting* e *underfitting* para os classificadores desse estudo. *Overfitting* indica que o modelo se adaptou demais ao conjunto de treinamento, mas não consegue generalizar o comportamento para exemplos não vistos. Esse fenômeno acontece em todas as aparições do Random Forest, por exemplo. Já o *underfitting* mostra que o modelo treinado não conseguiu identificar os padrões e ruídos nem no conjunto de treinamento, obtendo resultados ruins para as etapas de treinamento e de teste. Tentou-se diversas variações distintas de parâmetros para os classificadores do estudo, mas nenhuma conseguiu solucionar os problemas encontrados. O *underfitting* pode ser visto para os classificadores MLP, Logistic Regression e SVC usando a técnica de Word Embedding, por exemplo.

Classif.	WE (fastText)		BoW		TF-IDF	
	treino	teste	treino	teste	treino	teste
RF	98.97%	25.64%	91.58%	25.75%	90.68%	25.02%
kNN	100.00%	24.30%	36.15%	23.38%	38.80%	26.98%
PA	56.49%	25.65%	99.97%	22.86%	99.97%	25.54%
NB	0.00%	0.00%	80.90%	30.89%	50.72%	29.25%
GB	75.28%	30.49%	92.04%	26.98%	97.79%	26.57%
XGB	0.00%	0.00%	42.17%	27.91%	45.96%	30.48%
MLP	24.64%	24.1%	98.89%	27.70%	49.90%	28.22%
LR	54.45%	31.82%	97.81%	30.48%	64.21%	31.20%
SVC	40.73%	32.13%	89.03%	30.27%	31.15%	29.35%

Tabela 1. Acurácias obtidas para a classificação multiclasse dos tweets obtidos a partir da etapa de expansão.

A Tabela 2 mostra os resultados obtidos para o segundo dataset com as oito emoções, que foi obtido na etapa de refinamento de tweets da Seção 3.2. Ao analisar a tabela, pode-se observar que os mesmos fenômenos de *overfitting* e *underfitting* que ocorreram para o primeiro dataset se repetem. A principal diferença é que os resultados do segundo dataset são ligeiramente inferiores, atingindo acurácias de até 22%, enquanto os resultados do primeiro conjunto de dados alcançam acurácias acima de 30%. Observa-se que o problema de multiclasse é bem mais difícil de ser resolvido do que o de classificação binária. Assim, os próximos experimentos modelam o problema como um de classificação binária.

A Tabela 3 explicita as melhores acurácias obtidas por todos os classificadores e combinações de emoções do estudo com o dataset obtido na etapa de expansão dos dados discutida na Seção 3.2. Os resultados obtidos foram bastante satisfatórios para algumas combinações de emoções. A combinação de Alegria/Tristeza gerou acurácias acima de 70%, indicando que o classificador consegue diferenciar bem as duas emoções. Também pode-se observar que as melhores acurácias obtidas possuem sentimentos bem contrastantes entre si, como Alegria/Tristeza e Raiva/Alegria. Percebeu-se também que

Classif.	WE(fastText)		BoW		TF-IDF	
	treino	teste	treino	teste	treino	teste
RF	86.78%	15.66%	80.66%	14.97%	82.17%	13.73%
kNN	87.39%	13.73%	28.92%	14.56%	33.83%	16.48%
PA	49.04%	19.64%	87.08%	17.85%	87.39%	17.44%
NB	0.00%	0.00%	79.43%	20.46%	67.68%	19.91%
GB	71.36%	17.30%	83.14%	16.20%	87.05%	17.99%
XGB	0.00%	0.00%	40.07%	18.68%	46.12%	17.72%
MLP	16.65%	16.62%	86.54%	20.74%	16.07%	16.21%
LR	55.60%	22.12%	86.26%	21.70%	70.74%	21.43%
SVC	33.17%	20.87%	82.31%	20.33%	17.20%	15.80%

Tabela 2. Acurácias obtidas para a classificação multiclasse dos tweets obtidos na etapa de refinamento.

Sentimentos	Acurácia de teste	Classificador
Raiva e Alegria	81,57%	BERT
Desgosto e Alegria	72,24%	BERT
Raiva e Ansiedade	71,61%	BERT
Alegria e Tristeza	70,17%	BERT
Raiva e Tristeza	69,60%	BERT
Ansiedade e Desgosto	69,11%	BERT
Desgosto e Confiança	67,78%	BERT

Tabela 3. Melhores acurácias para cada par de emoções, obtidas pelos classificadores binários no conjunto da etapa de expansão.

o classificador BERT sempre obteve a melhor performance nas combinações de emoções que apresentavam as acurácias mais altas do estudo.

A Tabela 4 exibe os melhores resultados para o conjunto de dados obtido na etapa de refinamento da rotulação dos tweets, explicada na Seção 3.2. É possível notar que ocorreram diferenças notáveis entre as acurácias obtidas entre cada um dos conjuntos de dados. No segundo dataset, as oito melhores combinações de emoções apresentaram acurácias de teste superiores a 70% enquanto, na primeira parte, apenas quatro resultados possuíram esse mesmo comportamento.

A maior acurácia obtida no segundo dataset foi ligeiramente inferior para a maior acurácia do primeiro dataset e, coincidentemente, refletem a mesma combinação de emoções: Raiva e Alegria. Isso indica que ambas as emoções são as mais divergentes entre si dentre todas as combinações possíveis e implicam em uma facilidade maior para os classificadores entenderem suas diferenças e aplicarem suas predições de forma mais coerente com a realidade. Também pode-se observar que a maior parte das combinações de emoções foram similares em ambos os conjuntos de dados, com exceção da combinação de Raiva e Tristeza que não apresentou resultados muito bons no segundo conjunto de dados. Esse comportamento era esperado, por serem emoções que podem estar presentes em uma mesma situação, gerando uma dificuldade maior na diferenciação das emoções, como observado no primeiro critério adotado do refinamento de tweets na Seção 3.2. Diferentemente dos resultados obtidos no primeiro conjunto de dados, existem casos em que o classificador Logistic Regression, utilizando a técnica de Word Embedding, acaba apresentando melhores resultados do que o BERT.

Sentimentos	Acurácia de teste	Classificador
Raiva e Alegria	78,27%	LR (WE)
Raiva e Confiança	78,09%	BERT
Alegria e Tristeza	76,80%	BERT
Desgosto e Alegria	75,19%	BERT
Ansiedade e Desgosto	72,40%	BERT
Alegria e Confiança	72,16%	LR (WE)
Desgosto e Confiança	70,67%	BERT
Raiva e Ansiedade	70,18%	LR (WE)

Tabela 4. Melhores acurácias para cada par de emoções, obtidas pelos classificadores binários no conjunto da etapa de expansão.

Data	Significado
05/03/2020	Primeiro caso de Covid no Rio de Janeiro
16/03/2020	Primeiro estado de emergência declarado
17/03/2020	Decreto de isolamento social
30/03/2020	Primeira prorrogação do isolamento social
13/04/2020	Segunda prorrogação do isolamento social
30/04/2020	Terceira prorrogação do isolamento social
08/05/2020	Quarta prorrogação do isolamento social

Tabela 5. Datas importantes do conjunto de dados.

4.2. Análise dos conjuntos de dados

Para entender as características do conjunto de tweets coletado para esse estudo, estabeleceram-se divisões distintas para agrupar o dataset em subconjuntos. A primeira divisão separou todos os tweets em doze semanas. A segunda divisão foi feita utilizando datas de acontecimentos importantes na pandemia para o estado do Rio de Janeiro. Foram levantadas sete datas distintas e, a partir dessas datas, tweets foram separados em dois subconjuntos distintos: dois dias anteriores ao acontecimento e dois dias após o acontecimento. Cada uma das datas levantadas encontra-se na Tabela 5. A partir de cada uma dessas divisões, foram feitos indicadores para visualizar as particularidades de seus dados.

Agrupamento por semanas

A Tabela 6 exibe bigramas e trigramas utilizados em cada semana desse conjunto de dados. É possível observar cinco assuntos predominantes: ceticismo com a Covid-19 e seus possíveis impactos na sociedade, preocupação com a pandemia, governo e suas ações durante a crise da Covid-19, informações gerais sobre os impactos da doença no Brasil e ações relacionadas ao isolamento social. O ceticismo aparece mais na primeira semana do estudo, com o uso de termos como "fake news", "nenhum caso" e "tá com medo". Nota-se que, ao decorrer das semanas, esse ceticismo desaparece, com exceção da Semana 3.

A preocupação com a Covid-19 e a pandemia pode ser visualizada em todas as semanas. Uma possível explicação é que essas menções poderiam estar ocorrendo devido a campanhas de conscientização no Twitter sobre os cuidados necessários contra a doença em uma época de muitas incertezas. Também existiu uma preocupação dos usuários com sua própria saúde e de outros através do uso de termos como "muita gente", "corro risco",

Subconjunto	Bigramas/Trigramas
Semana 1	passo álcool, fake news, nenhum caso, tá com medo, Brasil em 2020, álcool em gel
Semana 2	muita gente, corro risco, milhões de pessoas, Álcool em gel, PAROU o mundo, viaja e traz, acima de 60
Semana 3	casos confirmados, fake news, vou morrer, Muita gente, ambiente fechado, vc ficar, álcool em gel, M da Saúde, sair de casa, NAO E BRINCADEIRA
Semana 4	sendo fechadas, SACO CHEIO, vida normal, 15 dias, ficar em casa, Casos no Brasil, entra em colapso, cuidar das pessoas, álcool em gel, presidente da República, DECRETO RIO N°
Semana 5	15 dias, SECRETARIA MUNICIPAL, Taxa de letalidade, MUNICIPAL DE SAÚDE, Funcionários da Saúde, ficar em casa, higienização das Ruas, Tá muito difícil, Declaração de Emergência
Semana 6	atitude Irresponsável, hospital Icaraí, efeitos colaterais, vc se sair, dor de garganta
Semana 7	vou morrer, ninguém morra, levem a sério, ficar em casa, ficar 14 dias, trancada em casa, manter a empresa
Semana 8	muitas pessoas, 3 dias, vou morrer, dor de cabeça, calados na Crise, álcool em gel, o ÚNICO país
Semana 9	saúde pública, nenhum país, PRESIDENTE BOLSONARO, dor de cabeça, homem em casa
Semana 10	saúde pública, ngm respeita, fazendo churrasco, Ficar em casa, FICA EM CASA, tira a máscara, gripes e resfriados
Semana 11	saúde mental, pessoas morreram hoje, tossir na cara, pessoas com comorbidades
Semana 12	casos confirmados, gente morrendo, nao aguento, ficar em casa, sair de casa, medo de morrer, passadores de pano

Tabela 6. Bigramas e trigramas de maior frequência observados em tweets coletados de 01/03/2020 até 23/05/2020

“milhões de pessoas”, “vou morrer”, “ninguém morra”, “saúde mental”, etc. Esse comportamento pode ser observado em todas as semanas a partir da segunda, que foi a época em que os primeiros casos de Covid-19 surgiram no Brasil.

Menções aos governos federais e municipais e suas ações na pandemia podem ser observadas entre as Semanas 3 e 9. Há um maior foco nos governos municipais nas Semanas 4 a 6, onde é possível observar menções a órgãos como uma Secretaria Municipal de Saúde e “hospital Icaraí” e ações em cada cidade como “DECRETO RIO N°”, “higienização das Ruas” e “Declaração de Emergência”. Já o governo federal é citado esporadicamente durante as semanas como referências ao Ministério da Saúde e ao presidente Jair Bolsonaro.

Tweets contendo informações gerais sobre a pandemia parecem ter sido encontrados em algumas semanas, sem apresentar um padrão definido. Termos como “casos confirmados”, “Casos no Brasil”, “15 dias” e “Taxa de letalidade” tendem a indicar conteúdos informativos à população, sendo possivelmente feitos por usuários jornalísticos ou por usuários especializados na área médica.

Por fim, pode-se notar diversas menções a atitudes relacionadas ao isolamento social em grande parte das semanas a partir da Semana 2, com palavras favoráveis ao isolamento como “ficar em casa” e outras ocorrências desfavoráveis ao isolamento como “sair de casa”, “ngm respeita” e “fazendo churrasco”, que surgem a partir das últimas três semanas do dataset, o que indica uma certa perda de paciência com o isolamento social por parte de alguns usuários.

Subconjunto	Bigramas/Trigramas
Data 1 - Antes	peçoal viaja, fake news, tá com medo, álcool em gel
Data 1 - Depois	tá com medo, Brasil em 2020
Datas 2 e 3 - Antes	corro risco, milhões de pessoas
Datas 2 e 3 - Depois	casos confirmados, álcool em gel, hora de ficarmos, M da Saúde, culpa do BR, entrar em Pânico, garrafas de álcool, hospital de gente
Data 4 - Antes	Vitamina D, 11 letras, presidente da República, DECRETO RIO N°, a crise econômica, FICAR EM CASA
Data 4 - Depois	15 dias, SECRETARIA MUNICIPAL, sair de casa, MUNICIPAL DE SAÚDE, Taxa de letalidade
Data 5 - Antes	atitude irresponsável, ninguém morra, sociedade civil, sair a noite, hospital Icaraf, isolado n° 1, tô me sentindo, ditadura do Witzel
Data 5 - Depois	saco cheio, vou ficar, ficar em casa, vontade de sair
Data 6 - Antes	seguro desemprego, dor de cabeça, Fica em casa, monte de gente
Data 6 - Depois	91.589 casos, controle de acesso, sair de casa, resultado do exame
Data 7 - Antes	Vc sobreviver, casos confirmados, 1 saco, gripes e resfriados, roda de samba, ficar em casa, índice de letalidade, resolver minha vida,
Data 7 - Depois	fazendo churrasco, LOTAR HOSPITAL, muita gente, tira a máscara

Tabela 7. Bigramas e trigramas de maior frequência obtidos nos tweets antes e depois de cada data com algum acontecimento relevante

Agrupamento por datas importantes

Assim como feito no agrupamento por semanas, foram obtidos bigramas e trigramas utilizados em cada data levantada nesse agrupamento. A Tabela 7 representa esse resultado.

Nos períodos de tempo próximos à Data 1 (primeiro caso confirmado de Covid no RJ), a reação das pessoas estava tendendo a levantar dúvidas sobre a gravidade da doença e/ou se ela realmente tinha chegado no Brasil, como a presença de expressões como "peçoal viaja" e "tá com medo".

Já nas Datas 2 e 3, o comportamento das pessoas mudou drasticamente. Antes dessas duas datas, pouco se discutia sobre a pandemia da Covid-19, apresentando poucos bigramas e trigramas. Após essas duas datas, as pessoas começaram a ficar preocupadas com a doença e seus efeitos, podendo observar vários trigramas relacionados a isso como "hora de ficarmos", "entrar em Pânico" e "garrafas de álcool".

Nas datas seguintes, pôde-se notar alguns padrões: discussões sobre aspectos sanitários, políticos e econômicos relacionados à pandemia e reações contrárias ao isolamento social. É interessante notar que aparecem bigramas e trigramas relacionados a essas reações contrárias justamente após as Datas 4 a 7 que são prorrogações do isolamento social, indicando que as pessoas já estavam começando a ficar saturadas com a pandemia e desejando voltar às suas vidas normais. Por exemplo: "sair de casa" na Data 4, "vontade de sair" na Data 5, "sair de casa" na Data 6 e "tira a máscara" na Data 7.

4.3. Observações sobre a anotação manual de emoções

Ao longo da divulgação do formulário da Seção 3.2, muitas pessoas reportaram dificuldades na realização desse processo manual de emoções. Algumas dificuldades relatadas foram a dificuldade em diferenciar emoções e a dificuldade em escolher uma emoção em *tweets* contendo ironia ou sarcasmo. Para expressar essas dificuldades em números, foram identificados todos os *tweets* em que mais de uma emoção tivesse recebido o

maior número de anotações pelos usuários. Contabilizaram-se todas as combinações dessas emoções mais destacadas pelos usuários para que fosse possível observar quais combinações de emoções foram mais confundidas e difíceis de diferenciar pelos usuários da pesquisa. A contagem das combinações foi normalizada para facilitar a visualização dos dados e se encontra na Tabela 8.

	Raiva	Ansiedade	Desgosto	Medo	Alegria	Tristeza	Surpresa	Confiança
Raiva	0	0,4035087719	1	0,08771929825	0,1578947368	0,5438596491	0,1403508772	0,05263157895
Ansiedade	0	0	0,5789473684	0,5438596491	0,6140350877	0,8947368421	0,4210526316	0,5438596491
Desgosto	0	0	0	0,1754385965	0,1929824561	0,8771929825	0,4210526316	0,2631578947
Medo	0	0	0	0	0	0,5087719298	0,1403508772	0,05263157895
Alegria	0	0	0	0	0	0,1578947368	0,350877193	0,5263157895
Tristeza	0	0	0	0	0	0	0,3333333333	0,1228070175
Surpresa	0	0	0	0	0	0	0	0,1403508772
Confiança	0	0	0	0	0	0	0	0

Tabela 8: Contagem normalizada das emoções mais confundidas pelos participantes do processo de anotação manual de emoções.

A Tabela 8 possui valores destacados por cores que representam níveis de dificuldade de diferenciação para as emoções destacadas, onde verde é um nível baixo, amarelo representa um nível médio, laranja é um nível alto e vermelho representa um nível muito alto. As células com o fundo na cor branca devem ser desconsideradas da análise. Pode-se notar que a combinação das emoções de Raiva e Desgosto foi a mais difícil de ser diferenciada pelos participantes da pesquisa, sendo acompanhada de perto pelas combinações de Ansiedade/Tristeza e Desgosto/Tristeza. Coincidentemente, as combinações Raiva/Desgosto e Desgosto/Tristeza foram utilizadas como critério de desempate para a construção do segundo conjunto de dados, conforme relatado na Seção 3.2. Essas combinações apresentaram desempenho ruim nos classificadores do segundo dataset na análise de emoções desenvolvida na Seção 4.1, indicando uma certa dificuldade em diferenciar esses pares de emoções, tanto pelas pessoas quanto pelos classificadores.

5. Conclusão

Esse artigo apresentou uma metodologia para análise de emoções em tweets em português relacionados à Covid-19 publicados no Rio de Janeiro. As emoções compreendem oito tipos que estão relacionadas com a roda das emoções [Plutchik 1984]. Dois conjuntos de dados foram criados, seguindo um processo de rotulação manual realizado por voluntários, onde o segundo conjunto é um refinamento da rotulação do primeiro para aliviar a discordância entre os participantes. Dada a dificuldade de obter resultados satisfatórios em um problema multiclasse com as oito emoções, procedeu-se à criação de classificadores binários que pudessem distinguir um par de emoções apenas. Após a obtenção dos resultados com vários classificadores e métodos de extração de atributos textuais, é possível concluir que a decisão tomada sobre o foco maior na construção de classificadores binários mostrou-se acertada, deixando clara a diferença nos resultados entre as combinações binárias de emoções e a combinação de todas elas juntas.

Ao efetuar a análise dos dados com duas distribuições distintas, foi possível encontrar temas em comum e achar possíveis explicações para os comportamentos observados de acordo com as publicações dos usuários. O ceticismo de parte da população com a Covid-19 ocorreu em uma época próxima a divulgação do primeiro caso da doença no Brasil, quando a emergência sanitária no país e no estado do Rio de Janeiro ainda não haviam sido decretadas. Preocupações com a Covid-19 e seus impactos na sociedade podem

ser observados na maior parte do dataset e indicam que, desde que o estado de emergência foi decretado, há uma grande atenção por parte da sociedade nesse assunto, como era de se esperar. Por fim, reações favoráveis ao isolamento podem ser bastante vistas no conjunto de dados, enquanto o surgimento de reações desfavoráveis provavelmente está diretamente vinculado às diversas prorrogações do isolamento social, gerando reações de inquietação em diversos usuários do Twitter. Foi possível notar também que, no processo de rotulação de emoções, existiram emoções que possuíam similaridades entre si como Desgosto/Tristeza e Raiva/Desgosto, o que explica algumas dificuldades encontradas pelos participantes desse processo e também os resultados não tão bons obtidos com essas combinações na análise de emoções desenvolvida.

Como trabalhos futuros, uma possível análise de emoções mais robusta pode ser feita para todas as oito emoções juntas em um mesmo processo de classificação. Devido aos resultados ruins obtidos com essa divisão, esse estudo focou nos classificadores binários desenvolvidos. Um ponto de atenção é que esse problema de classificação multiclasse é razoavelmente mais complexo do que o problema binário mas, certamente, os resultados obtidos seriam interessantes. Ainda em relação ao processo de classificação, uma abordagem instigante seria analisar por quanto tempo sua taxa de acertos perdura. Caso o classificador comece a perder sua eficácia ao longo do tempo, poderiam ser criados mecanismos para atualização do modelo treinado, sem que fosse necessário proceder com todo o processo de rotulação e aprendizado do zero.

Referências

- Baziotis, C., Pelekis, N., and Doulkeridis, C. (2017). Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis. In *Proc. of the 11th Int. Workshop on Semantic Evaluation (SemEval-2017)*, pages 747–754, Vancouver, Canada. Association for Computational Linguistics.
- Brum, P. V., Teixeira, M. C., Miranda, R., Vimieiro, R., Jr, W. M., and Pappa, G. L. (2020). A characterization of portuguese tweets regarding the covid-19 pandemic. In *Anais do VIII Symposium on Knowledge Discovery, Mining and Learning*, pages 177–184, Porto Alegre, RS, Brasil. SBC.
- Devlin, J., Chang, M., Lee, K., and Toutanova, K. (2019). BERT: pre-training of deep bidirectional transformers for language understanding. In Burstein, J., Doran, C., and Solorio, T., editors, *Proc. of the 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. ACL.
- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3):146–162.
- Honnibal, M., Montani, I., Van Landeghem, S., and Boyd, A. (2020). spaCy: Industrial-strength Natural Language Processing in Python.
- Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2017). Bag of tricks for efficient text classification. In Lapata, M., Blunsom, P., and Koller, A., editors, *Proc. of the 15th Conf. of the European Chapter of the Association for Computational Linguistics, EACL 2017, Volume 2: Short Papers*, pages 427–431. Association for Computational Linguistics.

- Ko, Y. (2012). A study of term weighting schemes using class information for text classification. In *Proc. of the 35th Int. ACM SIGIR Conf. on Research and development in information retrieval*, pages 1029–1030.
- Li, I., Li, Y., Li, T., Alvarez-Napagao, S., Garcia-Gasulla, D., and Suzumura, T. (2020). What are we depressed about when we talk about covid-19: Mental health analysis on tweets using natural language processing. In *Int. Conf. on Innovative Techniques and Applications of Artificial Intelligence*, pages 358–370. Springer.
- Loria, S. (2018). textblob documentation. *Release 0.15*, 2.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. In Bengio, Y. and LeCun, Y., editors, *1st Int. Conf. on Learning Representations, ICLR 2013, Workshop Track Proceedings*.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. In Burges, C. J. C., Bottou, L., Ghahramani, Z., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119.
- Müller, M., Salathé, M., and Kummervold, P. E. (2020). Covid-twitter-bert: A natural language processing model to analyse covid-19 content on twitter.
- Nemes, L. and Kiss, A. (2021). Social media sentiment analysis based on covid-19. *Journal of Information and Telecommunication*, 5(1):1–15.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Plutchik, R. (1984). Emotions: A general psychoevolutionary theory. *Approaches to emotion*, 1984:197–219.
- Roelleke, T. and Wang, J. (2008). Tf-idf uncovered: a study of theories and probabilities. In *Proc. of the 31st annual Int. ACM SIGIR Conf. on Research and development in information retrieval*, pages 435–442.
- Rustam, F., Khalid, M., Aslam, W., Rupapara, V., Mehmood, A., and Choi, G. S. (2021). A performance comparison of supervised machine learning models for covid-19 tweets sentiment analysis. *PloS one*, 16(2):e0245909.
- Scott, S. and Matwin, S. (1999). Feature engineering for text classification. In *ICML*, volume 99, pages 379–388. Citeseer.
- Souza, F., Nogueira, R., and Lotufo, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In *9th Brazilian Conf. on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In Guyon, I., von Luxburg, U., Bengio, S., Wallach, H. M., Fergus, R., Vishwanathan, S. V. N., and Garnett, R., editors, *Advances in Neural Information Processing Systems 30: Annual Conf. on Neural Information Processing Systems 2017*, pages 5998–6008.