DOI: 10.5748/20CONTECSI/PSE/AIT/7238

eLocator: e207238

ANALYSIS OF SENTIMENTS IN SHORT TEXTS FROM EMPATHY GAMES
ANÁLISE DE SENTIMENTOS EM TEXTOS CURTOS DE JOGOS EMPÁTICOS

Ramon Neirão Mendes – https://orcid.org/0009-0006-3468-9709
Universidade Federal Do Pará

Syanne Karoline Moreira Tavares – https://orcid.org/0009-0009-4267-682X
Universidade Federal Do Pará

Douglas Maia Dos Santos – https://orcid.org/0000-0003-2880-0571
Universidade Federal Do Pará

Fabíola Pantoja Oliveira Araujo – https://orcid.org/0000-0002-8485-5799
Universidade Federal Do Pará

ANALYSIS OF SENTIMENTS IN SHORT TEXTS FROM EMPATHY GAMES

ABSTRACT

Sentiment analysis (SA) plays an essential role in computing, focusing on enabling machines to understand and interpret human language, helping to extract subjective information from texts and videos, such as feelings. Thus, the present article aims to demonstrate the way the test data was collected, the method of classifying the collected dataset, and evaluating the results obtained with the trained BERT (Bidirectional Encoder Representations from Transformers) model, which is a modern model widely used in Natural Language Processing (NLP) tasks, for sentiment analysis in short texts of empathetic games in Brazilian Portuguese. The main results show that among the analyzed sentiments, only the texts classified as neutral obtained a result below 75% in the AUC (area under the ROC curve) metric when predicting the sentiment. The other sentiments used achieved results above 80% using the same metric. Furthermore, the study explores possible improvements at the time of training in order to improve the model.

Keywords: Sentiment Analysis; Deep Learning; Natural Language Processing; BERT.

ANÁLISE DE SENTIMENTOS EM TEXTOS CURTOS DE JOGOS EMPÁTICOS

RESUMO

A análise de sentimentos (AS) desempenha um papel imprescindível na computação, que se concentra em permitir que as máquinas compreendam e interpretem a linguagem humana, ajudando a extrair informações subjetivas de textos e vídeos, como sentimentos. Assim, o presente artigo tem por objetivo apresentar a forma como foram coletados os dados de teste, o modo de classificação do conjunto de dados coletados e avaliar os resultados obtidos com o modelo treinado BERT (*Bidirectional Encoder Representations from Transformers*), que é um modelo de aprendizado profundo amplamente utilizado em tarefas de processamento de linguagem natural (PLN), na análise de sentimentos em textos curtos de jogos empáticos em português brasileiro. Os principais resultados obtidos demonstram que dentre os sentimentos analisados, somente os textos classificados como neutros obtiveram um resultado abaixo de 75% na métrica AUC (*Area Under the Curve*) no momento de predizer o sentimento, os outros sentimentos utilizados ficaram acima de 80% utilizando a mesma métrica. Além disso, o estudo explora possíveis melhorias no momento do treinamento a fim de aprimorar os resultados do modelo.

Palavras-chave: Análise de Sentimentos; Aprendizado Profundo; Processamento de Linguagem Natural; BERT

1. INTRODUÇÃO

A análise de sentimentos é uma tarefa fundamental que visa extrair, processar e classificar opiniões, sentimentos, emoções e avaliações a partir do vasto conteúdo disponível na Web (Bing Liu, 2012). Esse campo de estudo tem se tornado uma área de estudo em crescimento constante, impulsionada pelo aumento exponencial do uso de mídias sociais e aplicativos de mensagens instantâneas. Com o advento dessas plataformas, as pessoas têm encontrado maneiras mais rápidas e concisas de expressar suas ideias, especialmente devido às restrições de caracteres. Nesse contexto, compreender e interpretar as emoções e

sentimentos transmitidos por textos curtos tornou-se crucial para compreender a opinião pública sobre determinados produtos ou até mesmo monitorar o bem-estar emocional das pessoas.

Além disso, a análise de sentimentos também encontra aplicação no campo de jogos empáticos, uma categoria de jogos que busca criar experiências que promovam empatia e compreensão emocional (Kamila Rodrigues, 2021). Esses jogos apresentam narrativas profundas e envolventes, permitindo que os jogadores explorem diferentes perspectivas e vivenciem uma ampla gama de sentimentos. A análise de sentimentos em textos curtos, nesse contexto, desempenha um papel fundamental na compreensão emocional desses jogos e na experiência dos jogadores.

A análise de sentimentos em textos curtos é um desafio complexo devido à natureza concisa e muitas vezes ambígua dessas mensagens. O processamento de linguagem natural (PLN) tem desempenhado um papel fundamental nesse campo, fornecendo técnicas e algoritmos capazes de extrair informações semânticas e emocionais de textos breves. Com o auxílio de abordagens como aprendizado de máquina e mineração de dados, é possível identificar padrões linguísticos, reconhecer contexto e classificar sentimento expresso em um texto.

Nesse artigo, será explorada a importância da análise de sentimentos em textos curtos, com discussão sobre suas aplicações práticas no contexto dos jogos empáticos. Serão apresentadas as principais técnicas e abordagens utilizadas nesta área, incluindo uso de redes neurais profundas. Também discutiremos os desafios enfrentados nesse tipo de análise, como a ambiguidade textual, e as estratégias adotadas para superá-los.

Logo, a análise de sentimentos em textos curtos desempenha um papel crucial na compreensão das emoções humanas. Além disso, no contexto dos jogos empáticos, essa análise permite um melhor entendimento do impacto emocional dessas experiências. Através do desenvolvimento contínuo de técnicas e algoritmos, é possível extrair conteúdos valiosos dessas mensagens breves, ajudando a assimilar melhor os sentimentos coletivos, monitorar tendências e aprimorar as interações digitais e experiências de jogo.

2. METODOLOGIA

A metodologia adotada neste artigo segue algumas das etapas indicadas por Hammes e Freitas (2014), que incluem a apresentação dos datasets, detalhando sobre a construção do mesmo com auxílio de *Web Scraping*, a concepção do modelo, o processo de *fine-tuning* e a avaliação dos resultados.

2.1 Web Scraping

O *Web Scraping* se trata do processo de extrair automaticamente informações e dados de sites da internet por qualquer meio que não seja um programa interagindo com uma API (*Application Programming Interface*). Nesse sentido, no intuito de realizar a coleta massiva de comentários de jogos empáticos, foi desenvolvido um projeto em *Python* para obter automaticamente *reviews* da plataforma *Steam*, principal loja online de jogos. Nele, foram utilizados como base de construção do código o documentado por Andrew Muller (2021). Com a utilização das bibliotecas como *requests* e *BeautifilSoup* foram extraídos

260 comentários do jogo "Last Day of June", 1901 do "Life Is Strange", 25 do "That Dragon, Cancer" e 2 do "Fragments of Him", todos em português brasileiro.

Os passos realizados envolveram a coleta do número de identificação de cada jogo, e a partir disso foi realizada a coleta de dados com a biblioteca *requests* que permitiu o armazenamento das *reviews* em arquivos *JSON* contendo inúmeras informações como *id* de recomendação, a data de criação, de edição, e o comentário sobre o jogo. Com isso, foi levantada a necessidade de filtrar apenas os comentários e para isso foi criada uma função em *Python* para ler os arquivos *JSON* e armazenar cada comentário em uma lista. Por fim, com o auxílio da biblioteca *openpyxl*, foram criadas planilhas *XLSX* para registrar os comentários anteriormente listados. Dessa maneira foi possível coletar um total de 2188 comentários que contribuíram para construção dos *datasets* aqui utilizados.

Entretanto, durante o processo de limpeza dos dados foram retirados alguns comentários em inglês, comentários feitos por tópicos e também somente com *emojis*, pois estavam em desacordo com o objetivo do projeto, que era a análise de comentários em português brasileiro. Além disso, a exclusão desses comentários em inglês ajudou a manter a consistência dos dados e a evitar qualquer interferência na análise da empatia nos jogos, uma vez que os comentários em diferentes idiomas poderiam conter nuances e contextos distintos. A decisão de focar apenas nos comentários em português brasileiro permitiu uma análise mais precisa e relevante dos aspectos emocionais e empáticos relacionados aos jogos em questão.

2.2 Datasets

Os conjuntos de dados na análise de sentimentos são fundamentais para treinar e avaliar algoritmos e modelos de aprendizado de máquina com o objetivo de identificar a polaridade emocional em textos, como opiniões, avaliações, resenhas e *feedbacks* de clientes. Esses conjuntos de dados contêm amostras de texto anotadas manualmente com suas respectivas classificações de sentimento. Neste trabalho, foram utilizados 3 conjuntos de dados, a versão em português brasileiro do *dataset GoEmotions* (Hammes, 2021), o *TweetSentBR* (Brum, 2018) e o dataset coletado de comentários de jogos empáticos.

A versão utilizada do *dataset GoEmotions* possui 70000 sentenças manualmente anotadas retiradas do fórum *Reddit* na língua inglesa. Esse conjunto de dados possui 28 *labels*, sendo 27 emoções e uma label neutra (Demszky, 2020). Todas as sentenças em inglês foram traduzidas para português brasileiro utilizando o *Google translate* (Pereira, 2021). Como o objetivo deste artigo é realizar a análise de sentimentos em textos curtos no idioma português, foram reclassificadas as emoções nos sentimentos positivo, negativo e neutro, representado na Figura 1.

```python
def polarize_data(df):
    df.loc[df["admiration"] == 1 , 'polaridade'] = 2     # admiração
    df.loc[df["amusement"] == 1 , 'polaridade'] = 2      # divertimento
    df.loc[df["approval"] == 1 , 'polaridade'] = 2       # aprovação
    df.loc[df["caring"] == 1 , 'polaridade'] = 2         # cuidadoso
    df.loc[df["curiosity"] == 1 , 'polaridade'] = 2      # curiosidade
    df.loc[df["excitement"] == 1 , 'polaridade'] = 2     # excitação
    df.loc[df["gratitude"] == 1 , 'polaridade'] = 2      # gratidão
    df.loc[df["joy"] == 1 , 'polaridade'] = 2            # alegria
    df.loc[df["love"] == 1 , 'polaridade'] = 2           # amor
    df.loc[df["optimism"] == 1 , 'polaridade'] = 2       # otimismo
    df.loc[df["pride"] == 1 , 'polaridade'] = 2          # orgulho
    df.loc[df["realization"] == 1 , 'polaridade'] = 2    # realização
    df.loc[df["relief"] == 1 , 'polaridade'] = 2         # alívio

    df.loc[df["anger"] == 1 , 'polaridade'] = 0          # raiva
    df.loc[df["annoyance"] == 1 , 'polaridade'] = 0      # aborrecimento
    df.loc[df["confusion"] == 1 , 'polaridade'] = 0      # confusão
    df.loc[df["disappointment"] == 1 , 'polaridade'] = 0 # desapontamento
    df.loc[df["disapproval"] == 1 , 'polaridade'] = 0    # desaprovação
    df.loc[df["disgust"] == 1 , 'polaridade'] = 0        # desgosto
    df.loc[df["embarrassment"] == 1 , 'polaridade'] = 0  # constrangimento
    df.loc[df["fear"] == 1 , 'polaridade'] = 0           # medo
    df.loc[df["grief"] == 1 , 'polaridade'] = 0          # pesar
    df.loc[df["remorse"] == 1 , 'polaridade'] = 0        # remorso
    df.loc[df["sadness"] == 1 , 'polaridade'] = 0        # tristeza
    df.loc[df["nervousness"] == 1 , 'polaridade'] = 0    # nervosismo

    df.loc[df["neutral"] == 1 , 'polaridade'] = 1
    df.dropna(inplace=True)
    df["polaridade"] = df["polaridade"].astype(int)
    return df
```

Figura 1. Código usado para a reclassificação das emoções em polaridades, que cria uma nova coluna chamada polaridade que contém os valores 0, 1 e 2 representando os sentimentos (Fonte: Autor).

Após a reclassificação de alguma das emoções em sentimentos, uma nova coluna foi criada com os valores 0, 1 e 2, que correspondem aos sentimentos negativo, neutro e positivo, respectivamente. Com essa nova coluna, foram identificados 28.487 comentários com sentimento positivo, 18.423 comentários com sentimento neutro e 20.113 comentários com sentimento negativo, totalizando 67.023 sentenças polarizadas, representado na Figura 2.

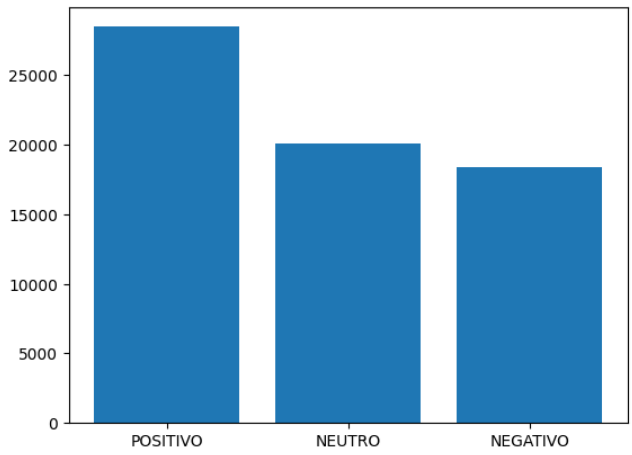

Figura 2. Gráfico da distribuição das polaridades do GoEmotions (Fonte: Autor).

O *TweetSentBR* é um *corpus* extraído manualmente do *Twitter*, contendo sentenças retiradas de hashtags (#) de programas de TV brasileiros durante a exibição dos mesmos (Brum, 2018). Esse conjunto de dados também é dividido em polaridades, contendo 3033 *tweets* classificados. Esses *tweets* são divididos em três partes: 1838 para treino, 323 para validação e 869 para teste, representadas na Figura 3.

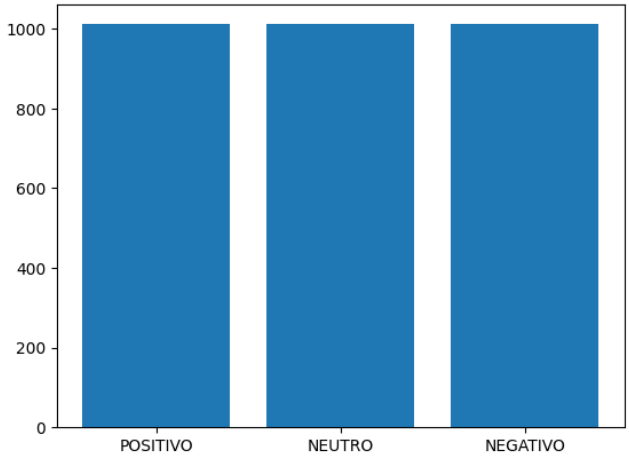

Figura 3. Gráfico da distribuição da polaridade do TweetSentBR (Fonte: Autor).

O dataset das avaliações de jogos empáticos foi retirado automaticamente por meio de *Web Scraping*, já descrito anteriormente, realizado nas lojas virtuais de jogos, como *Steam* e *Epic Games*, das análises dos usuários em português brasileiro. Foram escolhidos os jogos que atendessem a algumas das diretrizes de jogos empáticos, destacadas por Rodrigues (2021). A partir da coleta automática das avaliações foi realizada a classificação dos sentimentos manualmente em cada uma das sentenças coletadas, sendo 1488 positivos, 360 neutros e 221 negativos, totalizando 2059 comentários, representados na Figura 4.

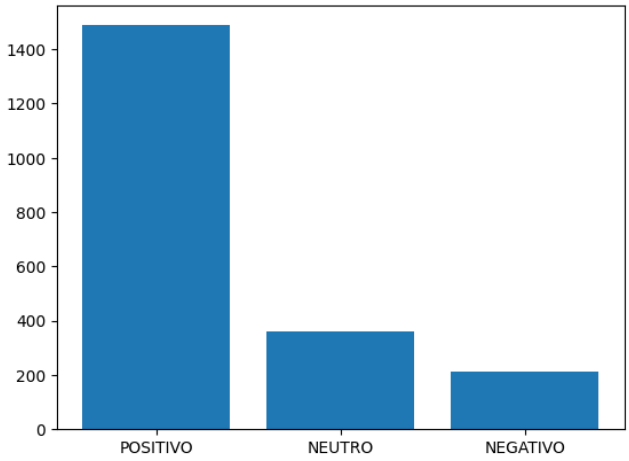

Figura 4. Gráfico de distribuição das polaridades dos jogos empáticos (Fonte: Autor).

2.3 Modelo

O modelo empregado foi o BERT (*Bidirectional Encoder Representations from Transformers*), é um modelo baseado na arquitetura *transformers*, que é um tipo de arquitetura de redes neurais, amplamente popular e influente no campo de PLN, introduzido por Vaswani (2017). A principal inovação dessa arquitetura é permitir que o modelo direcione sua atenção para partes específicas do texto de entrada de acordo com sua relevância para a tarefa. Uma das principais características do BERT é sua capacidade de leitura bidirecional, o que significa que ele pode compreender o contexto tanto das palavras anteriores quanto das palavras seguintes em uma sequência. Isso permite que o

modelo capture de forma mais precisa as nuances entre as palavras em um texto. No treinamento do BERT para a classificação de sentimentos em português brasileiro, foram utilizados os conjuntos de dados *GoEmotions* em conjunto com *TweetSentBR* (*train*), sendo testado com o conjunto de comentários de jogos empáticos coletados.

2.4 *Fine-tuning*

Fine-tuning é uma técnica de aprendizagem por transferência (*transfer learning*) amplamente empregada no campo de aprendizado de máquina e redes neurais profundas (*deep learning*), que permite ajustar um modelo pré-treinado em um conjunto de dados. No contexto deste trabalho foi realizado o *fine-tuning* do modelo BERT treinado para a classificação de sentimentos, os hiperparâmetros utilizados nesse processo estão detalhados na Tabela 1.

Hiperparâmetros	BERT
epochs	4
batch size	16
maximum sequence length	256
optimizer	*Adam*
learning rate	1e-5

Tabela 1. Hiperparâmetros utilizados no *fine-tuning*.

3. RESULTADOS

As métricas selecionadas para avaliar o desempenho do modelo foram a precisão (P), sensibilidade (*recall* ou S) e a medida-F (F1), sendo essa a média harmônica entre a precisão e sensibilidade, mesmas métricas utilizadas pelos autores Dorottya Demszky e Dana Movshovitz-Attias (2020). Também foram utilizadas as métricas, acurácia (A) e curva ROC (*Receiver Operating Characteristic Curve*) para uma melhor visualização de como cada sentimento está sendo classificado. Ao utilizar essas métricas, buscou-se avaliar tanto a precisão geral do modelo quanto sua capacidade de identificar corretamente os diferentes sentimentos presentes nos textos analisados. A acurácia mede a proporção de predições corretas em relação ao total de predições realizadas, fornece uma medida geral do desempenho do modelo, indicando a taxa de acertos em todas as classes, a precisão é a proporção de verdadeiros positivos em relação à soma de verdadeiros positivos e falsos positivos, é uma métrica útil quando é importante minimizar os falsos positivos, ou seja, reduzir os casos em que o modelo classifica erroneamente amostras negativas como positivas, a sensibilidade mede a proporção de verdadeiros positivos em relação à soma de verdadeiros positivos e falsos negativos, é especialmente importante quando o objetivo é identificar corretamente todas as amostras positivas, minimizando os falsos negativos e a medida-F é uma métrica que combina a precisão e a sensibilidade em uma única medida.

Pode-se observar na Tabela 2 que todas as métricas utilizadas ficaram em uma média de 70% a 76% , sendo as métricas coletadas do *dataset* de jogos empáticos ficou cerca de 3% acima em cada métrica, possivelmente devido a quantidade de positivos registrados.

	TweetSentBr (*Test*)	Jogos Empáticos
Acurácia	**0.75528**	**0.74439**
Precisão	**0.72985**	**0.76497**
Sensibilidade	**0.70804**	**0.74439**
Métrica-F	**0.71878**	**0.75454**

Tabela 2. Comparativo entre os testes feitos com os datasets TweetSentBR(*Test*) e o Jogos Empáticos.

A seguir serão analisadas as Curvas ROC, que possui dois parâmetros, a taxa de verdadeiros positivos (TVP) e a taxa de falsos positivos (TFP), calculados pelas quantidades de verdadeiros positivos (VP) quando o modelo diz que a classe é positiva e a classe era realmente positiva, verdadeiros negativos (VN) quando o modelo diz que a classe é negativa e a classe era realmente negativa, falsos positivos (FP) quando o modelo diz que a classe é positiva mas a classe era negativa e falsos negativos (FN) quando o modelo diz que a classe é negativa mas na realidade a classe era positiva.

A TVP é determinada pela quantidade de verdadeiros positivos sobre o total de casos verdadeiramente positivos, expressa pela fórmula:

$$TVP = \frac{VP}{(VP + FN)} \quad (1)$$

A taxa de falsos positivos é determinada pela quantidade de falsos positivos em relação ao total de casos que são verdadeiramente negativos. Essa métrica é expressa pela fórmula:

$$TFP = \frac{FP}{(FP + VN)} \quad (2)$$

Onde, os verdadeiros positivos em conjunto com os falsos negativos são o total de casos verdadeiramente positivos e os falsos positivos em conjunto com os verdadeiros negativos são os casos verdadeiramente negativos.

Uma curva ROC traça a taxa de verdadeiros positivos vs. taxa de falso positivos. Assim, na tentativa de simplificar a análise da ROC, a AUC (*Area Under the Curve*) nada mais é que uma maneira de resumir a curva ROC em um único valor, agregando todos os limiares da ROC, calculando a área sob a curva que varia de 1 a 0. Um modelo perfeito teria a área igual a 1.

Ao analisar as curvas ROC dos testes realizados (Figuras 5 e 6), nota-se uma significativa diferença, especialmente em relação ao sentimento neutro. Nos testes, as sentenças positivas e negativas apresentaram métricas AUC mais elevadas. Focando nos testes realizados com os comentários de jogos empáticos, os valores da métrica AUC foram de 0.86 para sentenças negativas e 0.80 para as positivas. No entanto, ao observar o valor das sentenças neutras, esse número caiu para 0.74 devido à forma como os comentários neutros foram classificados.

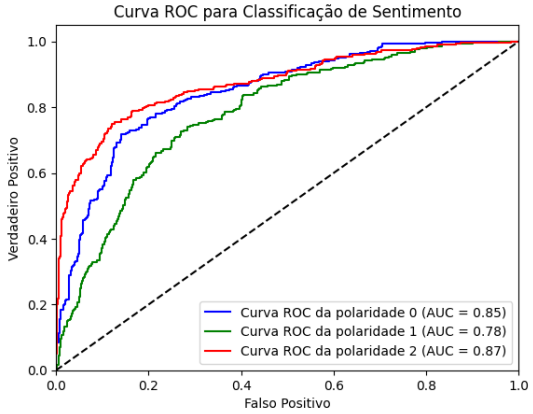

Figura 5. Curva ROC/AUC do TweetSentBR(*Test*) dividido por sentimento (Fonte: Autor).

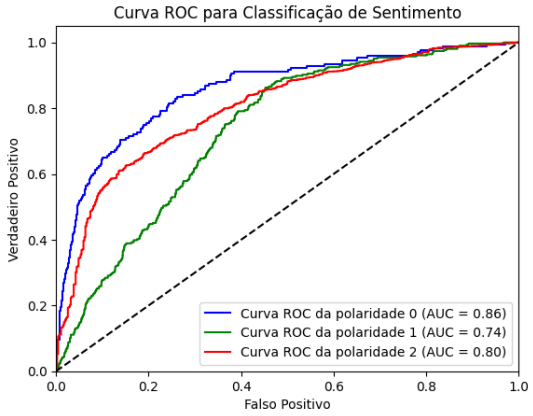

Figura 6. Curva ROC/AUC dos Jogos Empáticos dividido por sentimento (Fonte: Autor).

A classificação manual das sentenças neutras foi realizada a partir de comentários que não continham palavras ou expressões positivas ou negativas. Além disso, frases que continham expressões positivas e negativas em igual medida, sem a predominância clara de um sobre o outro, também foram classificadas como neutras. Por exemplo, frases como “Tão bonito quanto triste", “Que jogo lindo... e triste” e “Que jogo maravilhoso e triste ao mesmo tempo SOS” foram classificadas como neutras devido a falta de predominância de sentimento. Portanto, é possível afirmar que essa ambiguidade de alguns comentários resultou no valor mais baixo da métrica AUC para as sentenças neutras no modelo treinado.

Em seguida, serão examinadas as matrizes de confusão, que oferecem uma representação tabular do desempenho do modelo com a contagem de previsões corretas e incorretas de cada classe, em que cada linha representa a instância de uma classe predita, enquanto a coluna representa a instância da classe atual, permitindo uma análise mais precisa da classificação de cada sentimento.

Ao verificar as Figuras 7 e 8 é possível ver a mesma irregularidade que ocorre ao analisar as curvas ROC, dentro do escopo de sentimentos utilizados, no momento de conjecturar os sentimentos pelo modelo treinado com os corpos de dados *GoEmotions* e *TweetSentBR*, os comentários neutros são os preditos com maior erro.

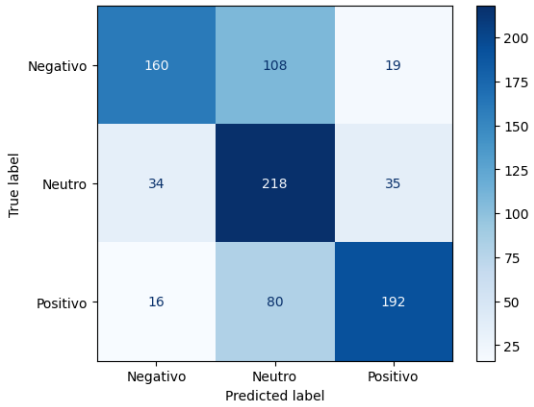

Figura 7. Matriz de confusão 3x3 da previsão dos sentimentos do TweetSentBR(*Test*) (Fonte: Autor).

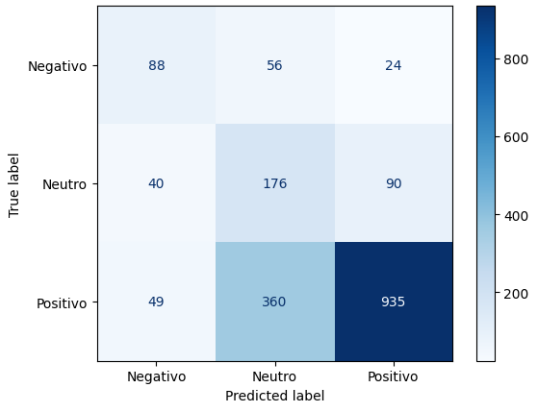

Figura 8. Matriz de confusão 3x3 da previsão dos sentimentos dos Jogos Empáticos (Fonte: Autor).

Para tentar minimizar o erro que ocorreu no período de predizer os sentimentos neutros, a medida utilizada foi colocar uma parte dos comentários coletados no treino, assim os exemplos ambíguos estarão no conjunto de treino tornando o modelo mais robusto e capaz de lidar com uma variedade de exemplos maior, inclusive aqueles que apresentam sentimentos neutros ambíguos. Com essa ação realizada, a Tabela 3 mostra como foram alterados os resultados em cada uma das métricas.

Após a abordagem implementada para aprimorar o desempenho do modelo, foram analisadas as métricas para avaliar sua eficácia e todas as métricas mostraram uma melhora significativa, apresentando ganhos médios de aproximadamente 10%, representados na Tabela 3.

Acurácia	**0.8476**
Precisão	**0.8569**
Sensibilidade	**0.8466**
Métrica-F	**0.8518**

Tabela 3. Resultado com os dados de Jogos Empáticos após a medida para o aprimoramento.

4. CONCLUSÃO E TRABALHOS FUTUROS

Por fim, ao avaliar todas as métricas obtidas, é possível concluir que a mineração de texto para a detecção e análise de sentimentos para jogos empáticos foi realizada com sucesso, mesmo que a detecção de comentários neutros esteja com a métrica AUC abaixo dos outros sentimentos. Isso ocorre devido a forma de classificação dos textos, podendo ser contornado com algumas mudanças no modo de treino do modelo, como por exemplo treinar com os comentários ambíguos.

Contudo, uma das dificuldades para a escrita do presente artigo foi a pequena quantidade de conjuntos de dados previamente classificados, disponíveis *on-line* para uso, e também a escassa quantidade de corpus em português brasileiro sem o uso de ferramentas de tradução instantânea.

Como próximos passos dessa pesquisa, está planejada a criação de uma plataforma de detecção das 6 emoções básicas descritas por Ekman (2003), incluindo felicidade, tristeza, raiva, surpresa, medo e nojo. Ainda como futuros objetivos, planeja-se treinar o modelo somente com frases coletadas totalmente em português brasileiro, uma vez que a maior parte do treinamento, realizado neste trabalho, foi efetuado com o *GoEmotions* que foi traduzido do inglês para português. Acredita-se que com essa prática, pode haver perda de contexto e tradução equivocada de algumas expressões idiomáticas.

5. REFERÊNCIAS

[1] Antunes , M. B. A. ., Issa , M. de F. ., & Hoed , R. M. . (2023). TÉCNICAS DE MACHINE LEARNING APLICADA A MINERAÇÃO DE DADOS E ANÁLISE DE SENTIMENTOS PARA PREDIÇÃO DE HOMOFOBIA NO TWITTER. REVISTA FOCO, 16(1), e853. https://doi.org/10.54751/revistafoco.v16n1-121.
[2] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5, 1--167. doi: 10.2200/s00416ed1v01y201204hlt016.
[3] Brum, H. B., & Nunes, M. das G. V. (2018). Building a sentiment corpus of tweets in brazilian portuguese. In Proceedings. Paris: European Language Resources Association. Recuperado de http://www.lrec-conf.org/proceedings/lrec2018/pdf/389.pdf.
[4]. Carvalho, C., Nagano, H., & Barros, A. (2017). A Comparative Study for Sentiment Analysis on Election Brazilian News. In Anais do XI Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana, (pp. 103-111). Porto Alegre: SBC.
[5] Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 4040–4054). Association for Computational Linguistics.
[6] Cambria, E. (2016). Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems, 31* (02), 102-107. https://doi.org/10.1109/MIS.2016.31.
[7] Souza, F., Nogueira, R., & Lotufo, R. (2020). 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* (pp. 403–417). Springer-Verlag, Berlin, Heidelberg. https://doi.org/10.1007/978-3-030-61377-8_28.
[8] Carvalho, F., Rodrigues, R., Santos, G., Cruz, P., Ferrari, L., & Guedes, G. (2019). Avaliação da versão em português do LIWC Lexicon 2015 com análise de sentimentos em

redes sociais. In Anais do VIII Brazilian Workshop on Social Network Analysis and Mining, (pp. 24-34). Porto Alegre: SBC. https://doi.org/10.5753/brasnam.2019.6545.
[9] Gomes, F. P., Silva, E. M., Teixeira, I., & Brito, P. F. (2017). Análise de Sentimentos: Uma Revisão Sistemática. In ENCOINFO - Congresso de Computação e Tecnologias da Informação, 19. (pp. 24–32). Palmas - TO: CEULP/ULBRA. ISSN e-ISSN: 2447-0767 (versão online). Disponível em: https://ulbra-to.br/encoinfo/edicoes/2017/artigos/analise-de-sentimentos-uma-revisao-sistematica/. Acesso em: 28 jul. 2023
[10] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics.
[11] Alammar, J. (2019). A Visual Guide to Using BERT for the First Time. Retrieved July 30, 2023, from http://jalammar.github.io/a-visual-guide-to-using-bert-for-the-first-time/
[12] Alammar, J. (2021). The Illustrated BERT, ELMo, and co. (How NLP Cracked Transfer Learning). Retrieved July 30, 2023, from https://jalammar.github.io/illustrated-bert/
[13] Rodrigues, K., Valério, C., M. Pinheiro, T., Maciel, C., & S. Nunes, E. (2021). Diretrizes para o Design de Jogos Empáticos. In Anais Estendidos do XX Simpósio Brasileiro de Jogos e Entretenimento Digital, (pp. 86-95). Porto Alegre: SBC. https://doi.org/10.5753/sbgames_estendido.2021.19628.
[14] Hammes, L., & Freitas, L. (2021). 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). Porto Alegre: SBC. DOI: https://doi.org/10.5753/stil.2021.17784.
[15] Muller, A. (2021). Scraping Steam User Reviews: Using Python to scrape user reviews from the Steam store without using the API. Retrieved July 30, 2023, from https://andrew-muller.medium.com/scraping-steam-user-reviews-9a43f9e38c92
[16] Ekman, P., & Friesen, W. V. (2003). *Unmasking the Face*. Cambridge: Malor Books.
[17] Pereira, D. A. (2021). A survey of sentiment analysis in the Portuguese language. *Artificial Intelligence Review*, 54 1087–1115. DOI: https://doi.org/10.1007/s10462-020-09870-1
[18] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is All You Need. In Advances in Neural Information Processing Systems (NIPS), 5998-6008. Disponível em: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf