

UTILIZANDO BERTIMBAU E TRADUÇÃO AUTOMÁTICA NA TAREFA DE CLASSIFICAÇÃO DE EMOÇÕES EM SENTENÇAS

LUIZ OTÁVIO ALVES HAMMES¹; LARISSA ASTROGILDO DE FREITAS²

¹Universidade Federal de Pelotas – loahammes@inf.ufpel.edu.br

²Universidade Federal de Pelotas – larissa@inf.ufpel.edu.br

1. INTRODUÇÃO

Análise de Sentimento (AS) é uma tarefa da área de Processamento de Língua Natural (PLN), que tem como objetivo extrair opiniões, sentimentos, emoções e avaliações da vasta quantidade de conteúdo disponível atualmente (LIU, 2012). A Classificação de Emoções (CE) em textos em nível de sentença é uma subárea da AS, a qual tem como objetivo classificar as distintas emoções expressas em um conjunto de sentenças que, em sua grande maioria, são de caráter subjetivo. Isso acontece pelo fato de que normalmente sentenças de caráter subjetivo são ricas em conteúdo sentimental, visto que, manifestam as percepções, os pontos de vista e o estado emocional dos seus autores (LIU, 2012).

De acordo com Pereira (2021), poucos são os trabalhos encontrados na literatura que focam na tarefa de CE em língua portuguesa. Dentre eles podemos citar o trabalho de Duarte *et al.* (2019) no qual foram utilizados *emojis* para reconhecer 6 tipos de emoções básicas (*happiness, anger, disgust, fear, sadness* e *surprise*) em *tweets*.

Neste trabalho, propomos utilizar modelos de Aprendizado de Máquina (AM) para classificação de 27 tipos de emoções com base no *dataset GoEmotions* (DEMSZKY *et al.*, 2020), traduzido para o português, com a finalidade de melhorar os resultados e ampliar a variedade de classes na tarefa de CE em língua portuguesa, disponíveis na literatura.

2. METODOLOGIA

O primeiro passo do trabalho foi realizar uma busca na literatura sobre *datasets* para CE em língua portuguesa. Não obtivemos sucesso na procura por *datasets* de tamanho razoável e com boa variedade de emoções anotadas no nível de sentença. Por esse motivo, em nossos experimentos utilizamos a versão filtrada do *dataset GoEmotions*, que contém 54263 sentenças manualmente anotadas em 28 classes (27 emoções + neutro) retiradas do fórum *Reddit* na língua inglesa. O processo de filtragem aplicado por Demszky *et al.* (2020) é realizado em duas etapas, primeiro são retirados todos os *labels* que foram selecionados por apenas um anotador e depois são mantidas apenas as sentenças com pelo menos um *label*. Manteve-se a proporção original: 80% treinamento, 10% validação e 10% teste.

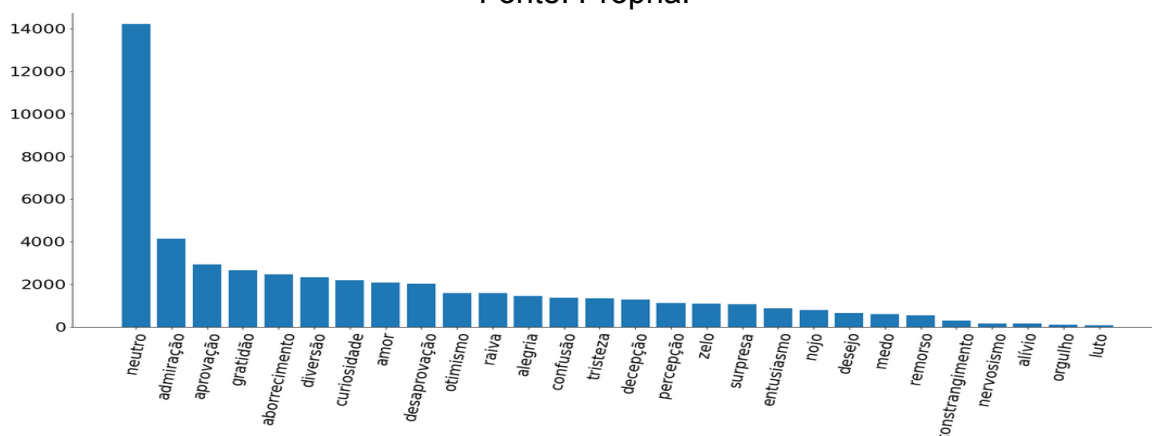
Como o objetivo deste trabalho é CE em sentenças escritas em português, traduzimos o *dataset GoEmotions* com o auxílio da biblioteca *itranslate* que facilita a utilização da API (*Application Programming Interface*) do *Google Translate*.

O modelo BERT (*Bidirectional Encoder Representations from Transformers*) (DEVILIN *et al.*, 2018), mais especificamente a versão pré-treinada para o português BERTimbau (SOUZA *et al.*, 2020), foi escolhido para esta tarefa, pois modelos baseados na arquitetura *Transformers* (VASWANI *et al.*, 2017), como o BERT, têm apresentado desempenho estado da arte nas mais diversas tarefas de PLN (GILLIOZ *et al.*, 2020).

Sobre a distribuição das quantidades de exemplos no *dataset GoEmotions* (Figura 1), é possível notar um grande desbalanceamento. A classe mais frequente “neutro”, tem aproximadamente 184 vezes mais amostras do que a classe “luto” que é a menos frequente. Caso nada for feito para amenizar este problema, o modelo classificador desenvolverá um viés, que resultará em uma habilidade preditiva ruim nas classes com poucos exemplos (ZHENG e JIN, 2020)

A fim de contornar esse obstáculo, foi utilizado o método *Class Balanced Loss* (CUI *et al.*, 2019). Neste método são calculados pesos para a função de *loss* do modelo utilizado, com base no número efetivo de amostras para cada classe. Um hiperparâmetro $\beta = 0.999$, também é utilizado para o cálculo destes pesos. Uma vez que, Cui *et al.* (2019) obteve um bom desempenho em boa parte dos seus experimentos utilizando este valor.

Figura 1: Distribuição das amostras por classe de emoção no *dataset GoEmotions*.
Fonte: Própria.



Fine-tuning é uma técnica de *transfer learning* na qual partindo de um modelo pré-treinado em uma tarefa de amplo domínio, os parâmetros desse modelo são ajustados para uma tarefa específica. Neste trabalho realizaremos o *fine-tuning* do BERTimbau-base para a tarefa de CE no nível de sentenças. Os hiperparâmetros utilizados no *fine-tuning* do modelo estão listados na Tabela 1.

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

Hiperparâmetros	
<i>epochs</i>	4
<i>batch size</i>	16
β (beta)	0.999
<i>maximum sequence length</i>	128
<i>warmup proportion</i>	0.2

3. RESULTADOS E DISCUSSÃO

Para avaliar o desempenho do modelo foram utilizadas as métricas Precisão, Revocação e Medida-F, mesmas métricas utilizadas pelos autores do *dataset GoEmotions* (DEMSZKY *et al.*, 2020), os quais utilizam em seus experimentos o modelo BERT-base para o inglês (DEVLIN *et al.*, 2018). A Tabela 2 apresenta a comparação entre os resultados reportados por DEMSZKY *et al.* (2020) e pela abordagem proposta, na tarefa de CE de 27 emoções no nível de sentença.

Podemos observar que ao utilizarmos a média macro da métrica Medida-F como parâmetro de avaliação do desempenho geral dos modelos analisados, o modelo para o português apresentou desempenho superior. É possível atribuir isso ao método de balanceamento utilizado (*Class Balanced Loss*), já que em outros experimentos realizados não se utilizou o método de balanceamento e o desempenho foi consideravelmente prejudicado.

Tabela 2: Comparativo entre os modelos para português e inglês.

Emoções	Modelo para o inglês BERT-base (DEVLIN et al., 2018)			Modelo para o português BERTimbau-base (SOUZA et al., 2020)		
	Precisão	Revocação	Medida-F	Precisão	Revocação	Medida-F
admiração	0.53	0.83	0.65	0.58	0.75	0.66
diversão	0.70	0.94	0.80	0.76	0.89	0.82
raiva	0.36	0.66	0.47	0.37	0.46	0.41
aborrecimento	0.24	0.63	0.34	0.37	0.33	0.35
aprovação	0.26	0.57	0.36	0.40	0.40	0.40
zelo	0.30	0.56	0.39	0.34	0.44	0.39
confusão	0.24	0.76	0.37	0.33	0.56	0.41
curiosidade	0.40	0.84	0.54	0.44	0.77	0.56
desejo	0.43	0.59	0.49	0.53	0.55	0.54
decepção	0.19	0.52	0.28	0.28	0.20	0.23
desaprovação	0.29	0.61	0.39	0.37	0.43	0.40
nojo	0.34	0.66	0.45	0.46	0.42	0.44
constrangimento	0.39	0.49	0.43	0.39	0.38	0.38
entusiasmo	0.26	0.52	0.34	0.36	0.43	0.39
medo	0.46	0.85	0.60	0.54	0.73	0.62
gratidão	0.79	0.95	0.86	0.88	0.91	0.90
luto	0.00	0.00	0.00	0.13	0.67	0.22
alegria	0.39	0.73	0.51	0.51	0.57	0.54
amor	0.68	0.92	0.78	0.72	0.87	0.79
nervosismo	0.28	0.48	0.35	0.29	0.48	0.36
neutro	0.56	0.84	0.68	0.64	0.70	0.67
otimismo	0.41	0.69	0.51	0.52	0.56	0.54
orgulho	0.67	0.25	0.36	0.39	0.44	0.41
percepção	0.16	0.29	0.21	0.37	0.12	0.18
alívio	0.50	0.09	0.15	0.14	0.27	0.18
remorso	0.53	0.88	0.66	0.51	0.88	0.64
tristeza	0.38	0.71	0.49	0.47	0.54	0.50
surpresa	0.40	0.66	0.50	0.49	0.60	0.54
média macro	0.40	0.63	0.46	0.45	0.55	0.48

Outro ponto importante que deve ser levado em consideração é a qualidade de tradução do *dataset*. Por exemplo, a frase: *“It’s a better option because it’s my life and none of your business? Lmfao, who are you”*, presente na base de treinamento, foi traduzida para: “É uma opção melhor porque é minha vida e nenhum da sua empresa? Lmfao, quem é você”, é possível perceber que na expressão idiomática *“none of your business”* foi realizada uma tradução literal dos seus termos, uma melhor tradução seria utilizar outra expressão idiomática com sentido equivalente para a língua portuguesa, como: “não é da sua conta”. Esse tipo de problema leva a uma deterioração, ou completa perda, do sentido completo expresso pelas

sentenças, podendo ser o suficiente para que a frase traduzida expresse uma emoção diferente, ou até mesmo, nenhuma emoção.

4. CONCLUSÕES

A partir dos problemas discutidos, podemos concluir que é necessária a criação de um *dataset* anotado, com uma boa variedade de emoções, no nível de sentença a ser utilizado na tarefa de CE em português, visto que, não encontramos esse tipo de recurso disponível na literatura. Ainda, percebemos que a utilização do método de balanceamento e o modelo em português obteve melhores resultados se compararmos com o modelo em inglês em um *dataset* traduzido automaticamente. Como trabalho futuro pretendemos investigar o uso de ferramentas de tradução automática capazes de identificar expressões idiomáticas e traduzí-las corretamente.

5. REFERÊNCIAS BIBLIOGRÁFICAS

LIU, B. Sentiment Analysis and Opinion Mining. **Morgan & Claypool Publishers**, 2012.

PEREIRA, D. A. A survey of sentiment analysis in the Portuguese language. **Artificial Intelligence Review**, v. 54, n. 2, p. 1087-1115, 2021.

DUARTE, L.; MACEDO, L.; OLIVEIRA, H. G. Exploring emojis for emotion recognition in portuguese text. In: **19th EPIA Conference on Artificial Intelligence**. Springer, p. 719-730, 2019.

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.

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.

SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: **9th Brazilian Conference on Intelligent Systems**. Springer, p. 403-417, 2020.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; ... POLOSUKHIN, I. Attention is all you need. In: **31st Advances in Neural Information Processing Systems**. p. 5998-6008, 2017.

GILLIOZ, A.; CASAS, J.; MUGELLINI, E.; ABOU KHALED, O. Overview of the Transformer-based Models for NLP Tasks. In: **15th Conference on Computer Science and Information Systems**. IEEE, p. 179-183, 2020.

ZHENG, W.; JIN, M. The effects of class imbalance and training data size on classifier learning: an empirical study. **SN Computer Science**, v. 1, n. 2, p. 1-13, 2020.

CUI, Y.; JIA, M.; LIN, T. Y.; SONG, Y.; BELONGIE, S. Class-balanced loss based on effective number of samples. In: **Conference on Computer Vision and Pattern Recognition**. IEEE, p. 9268-9277, 2019.