

Detecção de posicionamentos a partir de textos usando expansão semiautomática de córpus

Stance detection from text using semi-automatic corpus expansion

Camila Farias Pena Pereira ✉ 

Escola de Artes, Ciências e Humanidades - Universidade de São Paulo

Ivandr  Paraboni ✉ 

Escola de Artes, Ciências e Humanidades - Universidade de São Paulo

Resumo

A detecção computacional de posicionamentos—a tarefa de determinar, a partir de um texto de entrada, a atitude ou posição (e.g., favorável ou contrária) em relação a um t pico-alvo espec fico—geralmente depende de c rpus rotulados com informa  es de posicionamentos para cada t pico de interesse. Uma vez que esses t picos s o em princ pio ilimitados, a necessidade de novos c rpus rotulados tamb m o  . Como forma de amenizar algumas dessas dificuldades, este trabalho adapta para a detec  o de posicionamento um m todo de expans o de c rpus originalmente desenvolvido para a tarefa an loga de an lise de sentimento. O m todo   ent o aplicado a um grande (46 mil inst ncias) c rpus de posicionamentos cobrindo seis temas de interesse pol tico e/ou moral em portugu s brasileiro, obtendo um aumento substancial no n mero de inst ncias. Resultados de avalia  o autom tica e humana sugerem que a adi  o de inst ncias rotuladas semiautomaticamente ao conjunto de dados original n o prejudica a precis o da classifica  o, e que os r tulos gerados automaticamente est o, em sua maioria, corretos.

Palavras chave

detec  o de posicionamentos, expans o de c rpus

Abstract

Computational stance detection—the task of determining, given an input text, the attitude (e.g., for or against) towards a particular target topic—usually makes use of annotated corpus as training data and, since possible topics are in principle unlimited, so is the need for new labelled datasets about every topic of interest. In order to overcome some of these challenges, the present work adapts to the stance prediction task an existing corpus expansion method that has been originally devised for sentiment analysis. The method is applied to a large (46K instances) Brazilian Portuguese corpus conveying stances towards six target topics of moral and/or political nature, achieving a substantial increase in the number of labelled instances. Results from both automatic and human

evaluation suggest that adding semi-automatically labelled data to the corpus does not decrease accuracy, and that the majority of these labels are correct.

Keywords

stance detection, corpus expansion

1. Introdu  o

A detec  o computacional de posicionamentos (K    k & Can, 2020; Aldayel & Magdy, 2021) visa determinar, a partir de um texto fornecido como entrada, a atitude ou posi  o (por exemplo, favor vel ou contr ria) em rela  o a um determinado t pico-alvo de interesse. A tarefa permite identificar, por exemplo, se um indiv duo ou grupo est  concordando ou discordando de uma determinada afirma  o, se est  assumindo uma determinada posi  o sobre um t pico possivelmente controverso ou agressivo (da Silva et al., 2020) ou, de forma mais geral, estimar qual o reflexo de uma afirma  o presente no texto sobre o alvo (por exemplo, sendo a favor ou contra ele). Este  ltimo caso, tamb m conhecido como detec  o de posicionamentos orientada a alvos (do ingl s, *target-based stance detection*),   o foco do presente trabalho.

A detec  o de posicionamentos geralmente assume a forma de uma tarefa de aprendizado de m quina supervisionada baseada em c rpus anotados (por exemplo, postagens em redes sociais rotuladas manualmente com posi  es contr rias/favor veis a um alvo espec fico) e  , em princ pio, an loga   an lise de sentimento (AS), ou seja, a tarefa de determinar o sentimento positivo/negativo no texto (Zhang & v Wang, 2018).

No entanto, AS   uma tarefa de PLN possivelmente mais superficial no sentido de que os modelos de AS podem, em princ pio, usar qualquer



DOI: 10.21814/lm.16.2.436

This work is Licensed under a

Creative Commons Attribution 4.0 License

domínio suficientemente próximo (por exemplo, resenhas de filmes como em ‘*o filme foi terrível*’) como dados de treinamento para então inferir sentimentos em outros domínios (por exemplo, análises de produtos, como em ‘*a bateria do aparelho é péssima*’). A tarefa de detecção de posicionamentos, por outro lado, depende muito mais de conhecimento semântico associado ao tópico-alvo, fazendo-se necessário criar um novo conjunto de treinamento específico para cada alvo de interesse. Considere-se os seguintes exemplos:

- (1) *Claro que o certo é tomar cloroquina. Você ainda pode morrer de Covid-19, mas nunca de malária!*
- (2) *Se a vacina Coronavac é tão eficaz, porque é que nenhum país Europeu a utiliza?*

Ambos os exemplos expressam um posicionamento contrário a um medicamento ou tratamento que era discutido no contexto da pandemia de Covid-19. No entanto, além de misturar termos positivos (por exemplo, “certo”) e negativos (por exemplo, “morrer”), notamos que essas duas afirmações têm pouco em comum, por exemplo, em termos de vocabulário ou estrutura. Como resultado, um corpus de treinamento de posicionamento em relação a um tópico pode não ser necessariamente útil à construção de um modelo de detecção de posicionamento em relação a outro.

Uma possível estratégia para amenizar a escassez de dados rotulados seria a expansão de corpus de posicionamentos a partir de dados rotulados já existentes. De forma mais específica, o presente trabalho aborda uma estratégia de expansão de corpus originalmente desenvolvida para a tarefa de análise de sentimento por [Brum & das Graças Volpe Nunes \(2018\)](#), aqui adaptada para detecção de posicionamento na língua portuguesa. Esta estratégia foi aplicada a um experimento de expansão de corpus de posicionamentos, e resultados de avaliação intrínseca e humana são relatados.

O restante deste artigo está organizado da seguinte forma. Após uma visão geral dos trabalhos existentes na área de detecção de posicionamentos na seção 2, o presente trabalho é dividido em três partes principais. A seção 3 descreve o corpus de posicionamento a ser expandido automaticamente. A seção 4 apresenta os modelos classificadores a serem tomados como base para o método de expansão, seus resultados individuais e sua análise de características de aprendizado. Na seção 5 nossa atenção se volta para o método de expansão do corpus propriamente dito, descre-

vendo sua arquitetura e seus resultados. Por fim, a seção 6 apresenta nossas principais conclusões e sugestões de trabalhos futuros.

2. Trabalhos relacionados

A Tabela 1 apresenta uma visão geral dos estudos sobre a detecção computacional de posicionamentos, categorizados de acordo com o gênero de texto, idioma, modelo de representação textual e método computacional empregado. Detalhes adicionais são discutidos a seguir.

A partir deste levantamento, é possível constatar que a grande maioria dos estudos da área tende a ser baseada na rede social Twitter/X. Isso talvez seja explicável pela maior facilidade de coleta de dados que esta plataforma costumava oferecer. Além disso, assim como em outros tópicos de pesquisa em PLN, estudos de detecção de posicionamentos tendem a ser voltados ao idioma inglês, na maioria dos casos enfocando o uso de poucos conjuntos de dados tidos como padrão da área, como o corpus SemEval-2016 ([Mohammad et al., 2016](#)), VAST ([Allaway & McKeown, 2020](#)), e P-Stance ([Li et al., 2021](#)).

Com relação aos modelos de representação textual, é possível observar que os estudos da área seguem a evolução natural dos métodos empregados em outras áreas do PLN. Em estudos mais antigos, isso é sugerido pelo maior uso de modelos de contagens de *tokens* do tipo *bag-of-words* ou equivalente, frequentemente combinados com métodos de classificação tradicionais do tipo regressão logística ou SVMs. Ao longo dos anos, estas abordagens foram sendo substituídas pelo uso de *word embeddings* estáticos e posteriormente BERT, com uso crescente de métodos de aprendizado neural.

Em complemento ao levantamento das abordagens de detecção de posicionamentos existentes, foi realizado também um levantamento dos conjuntos de dados disponíveis considerando-se apenas a versão da tarefa orientada a alvos, que é o foco do presente trabalho. Os estudos deste tipo são sumarizados na Tabela 2.

3. O corpus UstanceBR r2

O presente estudo faz uso do corpus de posicionamentos *UstanceBR* versão r2, apresentado em [Pereira et al. \(2023\)](#). O corpus é uma coleção de *tweets* em língua portuguesa que foram manualmente anotados com informação referente ao posicionamento dos autores quanto a alvos de natureza altamente polarizada. Uma versão prévia

Estudo	Gênero	Idioma	Representação textual	Método
Zhou et al. (2019)	T	En	WE	CNN
Tun & Hninn Myint (2019)	T	En	BoW	NB
Siddiqua et al. (2019)	T	En	WE	LSTM
Sane et al. (2019)	T	Hi	BoW,WE	CNN
Li & Caragea (2019)	T	En	WE,l	LSTM
Fang et al. (2019)	C	En	BERT	TE
Mayfield & Black (2019)	D	En	WE,BERT	LR
dos Santos & Paraboni (2019)	T	Pt	BoW	NB, LR, RN
Xu et al. (2019a)	T	En	WE	RNN
Hassan & Lee (2019)	C	En	BoW	RN
Xu et al. (2019b)	D	En	WE	CNN,LSTM
Hanawa et al. (2019)	W	En	WE	LSTM
Bugueño & Mendoza (2019)	T	En	BERT	CNN
Schaefer & Stede (2019)	T	En	BoW,WE	SVM
Popat et al. (2019)	D	En	BERT	RN
D’Andrea et al. (2019)	D	It	BoW	SVM
Pavan et al. (2020)	T	Pt	WE	LSTM
Kovačević et al. (2020)	C	Se	l	-
Carvalho et al. (2020)	T	En	WE	LSTM
Conforti et al. (2021)	T,N	En	BERT	Softmax
Ayyub et al. (2021)	T	En	WE,PoS,l	NB,SVM,CNN
Su et al. (2021)	T	Ch	WE	Softmax
Samih & Darwish (2021)	T	En	BERT	SVM,Softmax
Zhu et al. (2021)	C	En	BERT	Softmax
Flores et al. (2022)	E	Pt	BERT	LSTM
Zhao et al. (2022)	T	En	BERT	Softmax
Pavan & Paraboni (2022)	T	Pt	BERT	LSTM
Ng & Carley (2022)	T	En	BERT	Softmax
Pavan et al. (2023)	C	Pt	BERT	LSTM
Chunling et al. (2023)	T	En	BERT	GAN
Zhang et al. (2023a)	T	En	none	ChatGPT
Wen & Hauptmann (2023)	N	En	BART	RN
Zhang et al. (2023b)	T,N	En	RoBERTa	RN
Cavalheiro et al. (2023)	T	Pt	BERT,BoW	BiLSTM,EN
Li et al. (2023)	T,N	En	BERT	ChatGPT
de Sousa & Becker (2023)	T	En	BoW	LR,SVM,RF
Zhang et al. (2024)	T,N	En	BERT	RN

Tabela 1: Estudos de detecção computacional de posicionamentos de acordo com gênero (A = avaliações; C = textos ou redações; D = debates Online; E = respostas à solicitações online; T = Twitter; N = notícias; ; W = Wikipedia, idioma (Ch = Chinês; En = Inglês; Hi = Hindi; It = Italiano; Pt = Português; Se = Sérvio), modelos de representação textual (BERT = BERT (Devlin et al., 2019); BoW = *bag-of-words*; D = Dependências Sintáticas; EL = características extralinguísticas; l = léxicos; PoS = *part-of-speech*; WE = Word Embeddings); e método computacional (CNN = Redes Neurais Convolucionais; EN = *ensemble*; GAN = redes neurais adversárias; ILP = *Inductive Logic Programming*; LSTM = *Long Short-term Memory networks*; RF = *Random Forest*; RL = Regressão Logística; NB = Naive Bayes; SVM = *Support Vector Machines*; RN = outros tipos de redes neurais).

deste corpus, chamada *UstanceBR* r1, foi apresentada em Pavan & Paraboni (2022) e não será aqui discutida.

O corpus *UstanceBR* r2 é composto de 86,8 mil *tweets*, totalizando cerca de 1,7 milhões de palavras, rotulados com informação de posicionamento binário (contra ou a favor) em relação a seis tópicos: presidentes brasileiros (Lula e Bolsonaro), medicamentos discutidos durante o período da pandemia Covid-19 (a vacina Corona-

vac e a Cloroquina), e instituições (a rede de TV Globo e a igreja). Os posicionamentos rotulados são aproximadamente balanceados por classe, e somam-se a um conjunto menor classificado como ‘Outros’ que contempla *tweets* para os quais não foi obtida concordância entre anotadores, ou que não apresentavam um posicionamento claro em relação ao alvo.

A Tabela 3 apresenta estatísticas descritivas do conjunto de posicionamentos rotulados do

Estudo	Gênero	Idioma	Instâncias (mil)	Rotulação
Mohammad et al. (2016)	T	En	4.2	t
Magdy et al. (2016)	T	En	336.3	p
Taulé et al. (2017)	T	Ca,Sp	10.8	t
Darwish et al. (2017)	T	Ar	33.0	t
Sobhani et al. (2017)	T	En	4.5	t
Conforti et al. (2020)	T	En	51.3	t
Pavan et al. (2023)	C	Pt	4.1	t,u
Mutlu et al. (2020)	T	En	14.4	t
Allaway & McKeown (2020)	O	En	23.6	t
Lai et al. (2020)	T	En,Fr,It,Sp,Ca	14.4	t
Glandt et al. (2021)	T	En	6.1	t
Jaziriyani et al. (2021)	T	Ar	9.6	t
Geiss et al. (2022)	R	En	2,717	u
Chen et al. (2022)	T	En,Fr,De,Du,Sp,*	17.9	t
Pereira et al. (2023)	T	Pt	86.8	t,u

Tabela 2: Córpus para detecção de posicionamentos orientada a alvo de acordo com gênero (T = Twitter ; R = Reddit ; C = textos ou redações ; O = opiniões), idioma (Ar=Árabe, Ca=Catalão, De=Alemão, Du=Holandês, En=Inglês, Fr=Francês, It=Italiano, Pt=Português, Sp=Espanhol, *=Outros), número de instâncias, e método de rotulação dos dados (t= em nível textual; p = propagação de rótulos ; u=em nível de usuário ou *timeline*).

córpus.

Para os experimentos descritos nas próximas seções, foram utilizados apenas os *tweets* com rótulos contra e a favor, com a divisão treino/teste denominada r2 em Pereira et al. (2023), que corresponde a 75% para treino e 25% para teste.

4. Detecção de posicionamentos

Como forma de avaliar o grau de dificuldade da tarefa de detecção de posicionamentos no córpus *UstanceBR*, e também estabelecer resultados de *baseline* a serem considerados na tarefa de expansão automática do córpus a ser discutida na próxima seção, foram conduzidos experimentos de classificação binária para os seis tópicos-alvo existentes.

Os experimentos conduzidos consideram o uso de três tipos de classificadores de posicionamentos que, dado um tópico-alvo de interesse, objetivavam atribuir um rótulo ‘a favor’ ou ‘contrário’ a cada *tweet* fornecido como entrada. Dois destes classificadores são do tipo regressão logística baseada em um modelo de representação textual *bag-of-words* utilizando contagem de palavras ou de caracteres, aqui denominados *RegLog.word* e *RegLog.char*, respectivamente. O terceiro classificador é baseado em um modelo pré-treinado de língua do tipo BERT (Devlin et al., 2019), que será aqui referido como *BERT*. O código destas implementações é disponibilizado publica-

mente para reuso¹.

Para os classificadores *RegLog.word* e *RegLog.char*, os seguintes parâmetros foram utilizados de forma padronizada para todos os tópicos-alvo: regularização com penalidade L_2 , tolerância para o critério de parada igual a $1e^{-1}$, número máximo de iterações igual a 100, e pesos das classes igual a 1. A modelagem de características é baseada na contagem Tf-Idf de palavras ou n-gramas de caracteres de tamanho 2 a 3, a depender do modelo em questão.

Para o classificador *BERT*, utilizou-se *BERTabaporu* (da Costa et al., 2023), um modelo BERT pré-treinado em um conjunto de 237 milhões de *tweets* em língua Portuguesa, entre os quais estão incluídos os textos de treino do córpus *UstanceBR*. Para treinamento e classificação, foi usada a classe *ClassificationModel* da biblioteca *Simple Transformers*² com as opções ‘*reprocess_input_data*’ e ‘*overwrite_output_dir*’ definidas como “true”. O treinamento foi feito em uma época, pois em testes preliminares não foi verificado um ganho de medida F_1 que justificasse o uso de um tempo maior. Para todos os demais parâmetros, foi usada a opção *default*, que inclui a definição de um *batch size* de 8 e o comprimento de sequência máximo de 128. A função *softmax* é acoplada à saída do

¹<https://github.com/camilafpp/expansaoUstanceBR>

²<https://simpletransformers.ai/docs/classification-models/>

Alvo	Contra	A favor	Outros	Total	Palavras	Palavras/ <i>Tweets</i>
Bolsonaro	5.565	3.848	1.668	11.082	259.521	23,42
Lula	4.514	3.806	13.934	22.254	422.064	18,97
Coronavac	4.058	3.915	932	8.905	252.663	28,37
Cloroquina	3.978	4.017	1.785	9.780	277.824	28,41
Globo	3.341	2.672	15.253	21.266	214.876	10,10
Igreja	3.539	3.598	6.329	13.466	322.289	23,93
Total	24.995	21.857	39.901	86.753	1.749.237	20,70

Tabela 3: Estatísticas descritivas do córpus UstanceBR r2 em Pereira et al. (2023).

Alvo	RegLog.word		RegLog.char		BERT	
	Acurácia	F_1	Acurácia	F_1	Acurácia	F_1
Bolsonaro	0,77	0,75	0,79	0,77	0,84	0,80
Lula	0,76	0,76	0,79	0,79	0,79	0,76
Coronavac	0,77	0,77	0,78	0,78	0,80	0,81
Cloroquina	0,80	0,80	0,80	0,80	0,80	0,80
Globo	0,79	0,78	0,81	0,81	0,85	0,81
Igreja	0,77	0,77	0,78	0,78	0,84	0,84

Tabela 4: Resultados da detecção de posicionamentos no córpus UstanceBR r2.

modelo para obtenção dos rótulos das predições.

4.1. Avaliação da tarefa de detecção de posicionamentos

Para fins de avaliação dos modelos descritos na seção anterior, foi utilizada a divisão padrão dos textos do córpus *UstanceBR* r2 (Pereira et al., 2023). Com base nesta divisão, os modelos propostos foram treinados e testados individualmente para cada um dos seis tópicos-alvo do córpus. Os resultados obtidos são sumarizados na Tabela 4, em que os melhores escores de medida F_1 para cada alvo estão destacados.

Os resultados apresentados na Tabela 4 indicam que o modelo *BERT* obteve os melhores resultados de classificação de forma geral, seguido pelo *RegLog.char*. Essa variação, no entanto, foi diferente para cada tópico, apresentando uma diferença de 7 pontos percentuais no caso do tópico ‘Igreja’ e de apenas um ponto no caso do alvo ‘Cloroquina’.

4.2. Cômputo da importância das características de aprendizado

Para melhor entender o comportamento dos modelos de classificação desenvolvidos, foi conduzida uma análise complementar dos termos considerados mais relevantes para cada classe, tomando-se por base para este fim o modelo *RegLog.word*. A escolha desse modelo para a análise foi motivada pela observação de que

RegLog.char e *BERT* são baseados em características de interpretação humana mais difícil.

A análise do modelo *RegLog.word* fez uso da biblioteca eli5³, que fornece uma explicação dos resultados do modelo com base nas suas características mais importantes, ou seja, calculando o peso relativo de cada *token* na classificação. Esses pesos, que são calculados a partir de permutações de características (Breiman, 2001), indicam a mudança no resultado da métrica de avaliação quando a característica em questão é substituída por outra. Os pesos assim computados são positivos para posicionamentos favoráveis e negativos para posicionamentos contrários, devendo-se considerar o seu valor absoluto para estimar sua importância. Os resultados obtidos são sumarizados na Tabela 5.

Com base nos resultados da Tabela 5, observa-se que as características mais importantes são, de forma geral, termos intuitivamente consistentes com os alvos dos posicionamentos, e não termos de pouco conteúdo semântico como pronomes ou preposições, o que poderia levantar suspeitas de *overfitting*. Nota-se também que as características mais importantes pertencem, de modo geral, a um vocabulário específico associado a cada tópico, e na maioria dos casos não correspondem a termos de caráter mais genérico que poderiam ser relacionados a mais de um posicionamentos ou tópico.

³https://eli5.readthedocs.io/en/latest/autodocs/eli5.html#eli5.explain_weights

Lula		Coronavac		Globo		Bolsonaro		Cloroquina		Igreja	
peso	caract.	peso	caract.	peso	caract.	peso	caract.	peso	caract.	peso	caract.
3,586	presidente	3,449	gado	3,464	amo	5,307	presidente	3,996	anos	2,989	vou
2,877	moro	3,351	bolsonaro	3,270	na	3,760	nosso	3,930	hidroxicloroq.	2,527	ir
2,820	contra	3,021	doses	2,725	parab�ns	3,360	m�dia	3,923	china	2,506	saudade
2,602	provas	2,769	butantan	2,507	s�rie	2,875	est�	3,805	vidas	2,409	nossa
2,509	golpe	2,352	coronavac	2,487	filme	2,741	imprensa	3,698	esquerda	2,259	amanh�
2,390	livre	2,165	bozo	2,369	obrigada	2,724	esquerda	3,425	governadores	2,185	hoje
2,308	cora��o	2,042	mil	2,251	novela	2,710	parab�ns	3,271	chin�s	1,999	maria
2,220	lula	1,984	vacinas	2,040	passando	2,642	povo	3,118	globo	1,952	fui
2,029	persegui��o	1,833	gente	2,008	plant�o	2,633	stf	3,101	azitromicina	1,857	indo
1,942	julgamento	1,822	instituto	1,771	bbb	2,595	apoio	2,803	uip	1,843	senhor
-4,582	bandido	-3,488	vachina	-5,170	lixo	-3,861	via	-3,561	gado	-3,571	universal
-4,227	corrup��o	-3,173	d�ria	-4,479	globolixo	-2,999	cu	-3,234	cloroquina	-3,107	evang�lica
-4,025	ladr�o	-2,852	contrato	-3,099	merda	-2,734	eu	-3,205	trump	-3,039	dinheiro
-3,409	condenado	-2,809	dinheiro	-2,247	seus	-2,566	merda	-2,772	bozo	-2,337	igrejas
-3,226	silva	-2,803	chinesa	-2,091	vergonha	-2,252	t�	-2,350	estoque	-2,314	eles
-3,088	cadeia	-2,658	doria	-2,084	tomar	-2,065	minha	-2,241	capit�o	-2,179	cat�lica
-3,061	presidi�rio	-2,574	china	-1,974	se	-2,000	coronav�rus	-2,170	oz�nio	-2,003	crente
-3,052	corrupto	-2,318	obrigar	-1,945	n�o	-1,968	ta	-2,119	gripezinha	-1,887	estado
-2,102	criminoso	-2,067	cobaia	-1,891	contra	-1,900	votou	-2,033	ah	-1,881	religi�o
-2,047	babaca	-1,965	essa	-1,782	bosta	-1,770	falando	-2,021	eua	-1,746	inferno

Tabela 5: Caracter sticas mais importantes no modelo *RegLog.word*.

Al m de os classificadores exibirem um comportamento razoavelmente pr ximo do esperado, observa-se ainda que a maioria das caracter sticas corresponde a palavras de conte do, como ‘presidente’, ‘doses’, ‘filme’, etc., e que n o podem ser classificadas como representando um sentimento positivo ou negativo na tarefa de an lise de sentimentos. Isso refor a o entendimento da natureza tradicional da tarefa de classifica  o de posicionamento como distinta da de classifica  o de sentimento, sendo possivelmente mais dependente de conhecimento sem ntico (K    k & Can, 2020).

5. Expans o de dados

Apesar dos resultados de classifica  o de modo geral positivos reportados na se  o anterior, observa-se que a massa de dados coletada nas etapas iniciais de constru  o do c rpus *UstanceBR* permanece em grande parte sem uso, dado que apenas uma pequena por  o (aproximadamente 24%) dos dados foi anotada por raz es de custo. A Tabela 6 apresenta estat sticas descritivas do conjunto de treino e testes rotulado manualmente, e o n mero de inst ncias n o rotuladas de cada subconjunto (alvo).

O aproveitamento de um volume maior de dados, desde que rotulados de forma confi vel, poderia viabilizar a constru  o de modelos mais robustos para a tarefa de classifica  o discutida na se  o anterior. Assim, consideramos

Alvo	Treino	Teste	N�o Rotulado
Bolsonaro	7.060	2.354	31.995
Lula	6.240	2.080	27.489
Coronavac	5.979	1.994	35.790
Cloroquina	5.996	1.999	35.181
Globo	4.509	1.504	29.854
Igreja	5.352	1.785	34.590

Tabela 6: Inst ncias de treino e teste rotulados manualmente e por  o n o rotulada do c rpus *UstanceBR* r2.

uma estrat gia de expans o semiautom tica do conjunto de dados existente usando os modelos *RegLog.word*, *RegLog.char* e *BERT* descritos na se  o anterior. O termo ‘semiautom tica’ denota aqui a combina  o do processo de anota  o manual inicial com a presente automa  o.

Para a expans o do c rpus, foi escolhido o m todo *self-training* de aprendizado semi-supervisionado (Zhu, 2005). Um procedimento similar foi utilizado em Brum & das Gra as Volpe Nunes (2018) para classifica  o de sentimento, utilizando o m todo de *self-training* para expandir um conjunto de dados de texto rotulado manualmente em um *framework* denominado CasSUL. Uma das premissas deste m todo   a de que tarefas com um certo grau de subjetividade, como seria o caso da pr pria an lise de sentimentos, n o podem prescindir de anota  o manual como uma parte do processo.

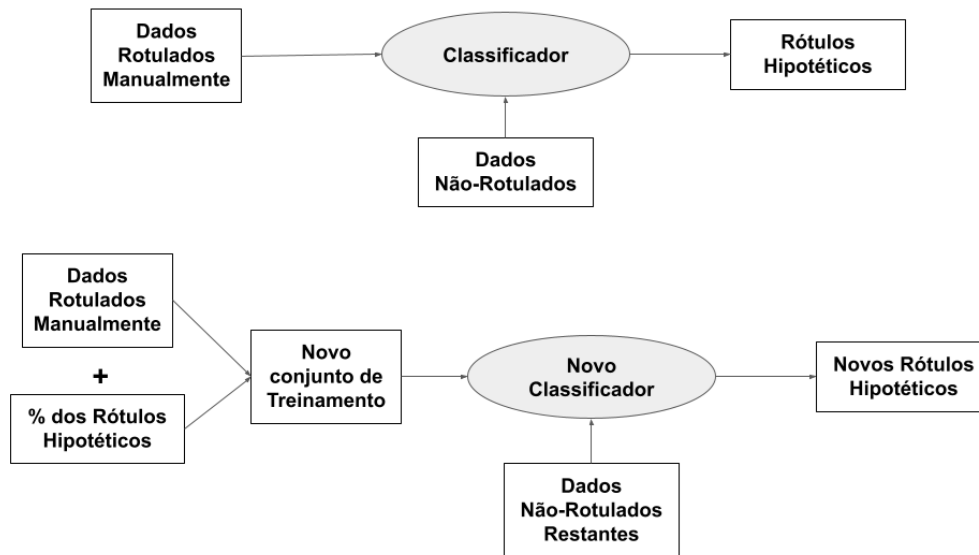


Figura 1: Modelo de expansão do conjunto de dados.

Em Brum & das Graças Volpe Nunes (2018), foram conduzidos experimentos utilizando seis classificadores: *Support Vector Machines* (SVM), *Bernoulli Naive Bayes*, Regressão Logística, *Multilayer Perceptron* (MLP), Árvores de Decisão e *Random Forest*. Para todas as alternativas exceto Regressão Logística, foi realizada uma busca em grade para definir os melhores parâmetros de cada modelo. Além disso, foram realizados experimentos com diversas combinações de características para além do texto puro, como uso de *emoticons*, *emoji* e *part-of-speech*. Os melhores resultados foram encontrados utilizando-se classificadores do tipo MLP e Regressão Logística, porém sem ultrapassar significativamente o resultado obtido com o conjunto rotulado manualmente. No entanto, o *framework* CasSUL permitiu expandir o conjunto de dados original em 8 vezes.

5.1. Procedimento de expansão

A Figura 1 apresenta uma representação gráfica do processo de expansão de córpus proposto. Na primeira parte, os dados anotados manualmente são utilizados para treinar um classificador utilizado para rotular dados de forma dita hipotética. A parte inferior ilustra como um percentual desses dados é adicionado ao conjunto rotulado manualmente, gerando um novo conjunto de treinamento. Um classificador é então novamente treinado com esses dados e o ciclo se repete.

De maneira análoga ao *framework* CasSUL, a expansão do conjunto de dados de posicionamentos também se inicia a partir dos modelos de classificação treinados com o conjunto rotulado manualmente. Esses modelos são então usados para classificar as instâncias restantes que, estando assim rotuladas de forma hipotética, são avaliadas para possível agregação ao córpus. Para esta avaliação de elegibilidade, as novas instâncias são ordenadas por confiabilidade, aqui definida como a probabilidade da predição fornecida pelo classificador. A qualidade do conjunto resultante após a adição de novos dados é testada treinando-se um novo modelo, e então avaliando-se seu desempenho com base na medida F_1 no mesmo conjunto de teste. O processo então se repete, com o novo classificador sendo usado para rotular os dados que não foram selecionados, até que não seja mais possível fazer novas adições dentro do limite dado pelo valor de *threshold* em uso.

Apesar de fazer uso do mesmo método adotado no *framework* CasSUL, o presente modelo apresenta algumas diferenças fundamentais em relação a Brum & das Graças Volpe Nunes (2018). Em primeiro lugar, observa-se que o presente trabalho trata de uma tarefa distinta da classificação de sentimento, que é a detecção de posicionamento em relação a um determinado alvo. Conforme destacado em Küçük & Can (2020), sentimento e posicionamento não estão necessariamente correlacionados. Para exemplificar, considere-se os dois textos abaixo, ambos expressando um posicionamento favorável ao alvo ‘Igreja’, porém com sentimentos opostos (ou seja, negativo no primeiro caso e positivo no segundo).

Alvo	RegLog.word		RegLog.char		BERT	
	Threshold	Iterações	Threshold	Iterações	Threshold	Iterações
Bolsonaro	1%	11	1%	16	1%	1
Lula	1%	13	10%	2	1%	1
Coronavac	1%	8	5%	2	1%	3
Cloroquina	1%	5	1%	2	1%	1
Globo	1%	6	1%	1	1%	3
Igreja	1%	14	5%	3	1%	5

Tabela 7: Parâmetros ótimos de cada modelo.

‘Eu vou na igreja para me proteger de todo esse mal.’

‘Fui na igreja hoje, me senti muito bem.’

Uma segunda diferença do presente estudo em relação ao de Brum & das Graças Volpe Nunes (2018) é a estratégia de modelagem de dados adotada, no presente caso utilizando não apenas a abordagem *bag-of-words* (BoW) tradicional, mas também o uso de *transformers* do tipo BERT (Devlin et al., 2019). Além disso, o presente trabalho enfoca apenas a construção de modelos puramente textuais, não contemplando o uso de características como *emoji*, *emoticons* e etiquetas de *part of speech* que foram utilizadas em Brum & das Graças Volpe Nunes (2018), e que apresentaram resultados inferiores ao de modelos textuais.

Finalmente, outra diferença significativa entre o presente trabalho e o *framework* CasSUL é ainda a questão do balanceamento dos rótulos gerados. Para evitar o problema de desbalanceamento de classes apontado em Brum & das Graças Volpe Nunes (2018), ao invés de fazer a seleção apenas com base na confiabilidade do classificador, no presente trabalho são selecionadas as N instâncias mais confiáveis de cada classe, onde N representa a metade do volume de dados a ser selecionado com base no *threshold* definido. Assim, o conjunto de dados permanece balanceado durante todo o procedimento, evitando que um desbalanceamento de classes viesse a ser propagado para as etapas seguintes de forma cumulativa.

Para determinar o número de instâncias a serem selecionadas, foram experimentadas diferentes proporções do conjunto não rotulado inicial considerando-se um valor de *threshold* de 1%, 5%, 10%, 25% e 40%. No caso do modelo BERT, entretanto, dado o custo computacional exigido para treino de múltiplas configurações de *threshold*, os experimentos foram limitados à adição de 1% do conjunto não-rotulado e até 5 iterações em todos os conjuntos.

A Tabela 7 sumariza os parâmetros ótimos de cada modelo referentes ao valor de *threshold* e número de interações empregados.

5.2. Resultados da expansão

A Figura 2 ilustra a evolução dos resultados do classificador *RegLog.word* com diferentes números de iterações e valores de *threshold*.

No caso do modelo *RegLog.word*, observa-se que a adição sucessiva de 1% dos dados rotulados alcançou o melhor desempenho, com ganho de até 3% na medida F_1 . Além disso, em todos os tópicos esse desempenho atinge seu máximo com a adição de menos de 20% dos dados não-rotulados, tendendo a diminuir com a incorporação de maiores volumes.

Já no caso do modelo *RegLog.char*, ilustrado na Figura 3, os resultados obtidos foram ligeiramente superiores. No entanto, a diferença entre o desempenho do modelo treinado com o conjunto de dados rotulado manualmente e o conjunto expandido foi modesta, com ganho máximo de 2% de medida F_1 .

Finalmente, a Figura 4 apresenta os resultados da expansão usando o modelo BERT com adição de uma porção fixa de 1% do conjunto não-rotulado, em que a interação ‘0’ diz respeito ao resultado obtido pelo conjunto rotulado manualmente. Como discutido, no caso deste modelo específico foi realizado um número menor de interações devido ao seu custo computacional, mas ainda assim foi possível observar um comportamento similar ao dos modelos *RegLog.word* e o *RegLog.char*. Este resultado sugere que não seriam obtidos resultados significativamente superiores aos atuais caso o procedimento fosse ampliado.

5.2.1. Avaliação Intrínseca

Como uma primeira forma de avaliar a qualidade dos rótulos atribuídos de forma semiautomática, foram conduzidos experimentos de classificação

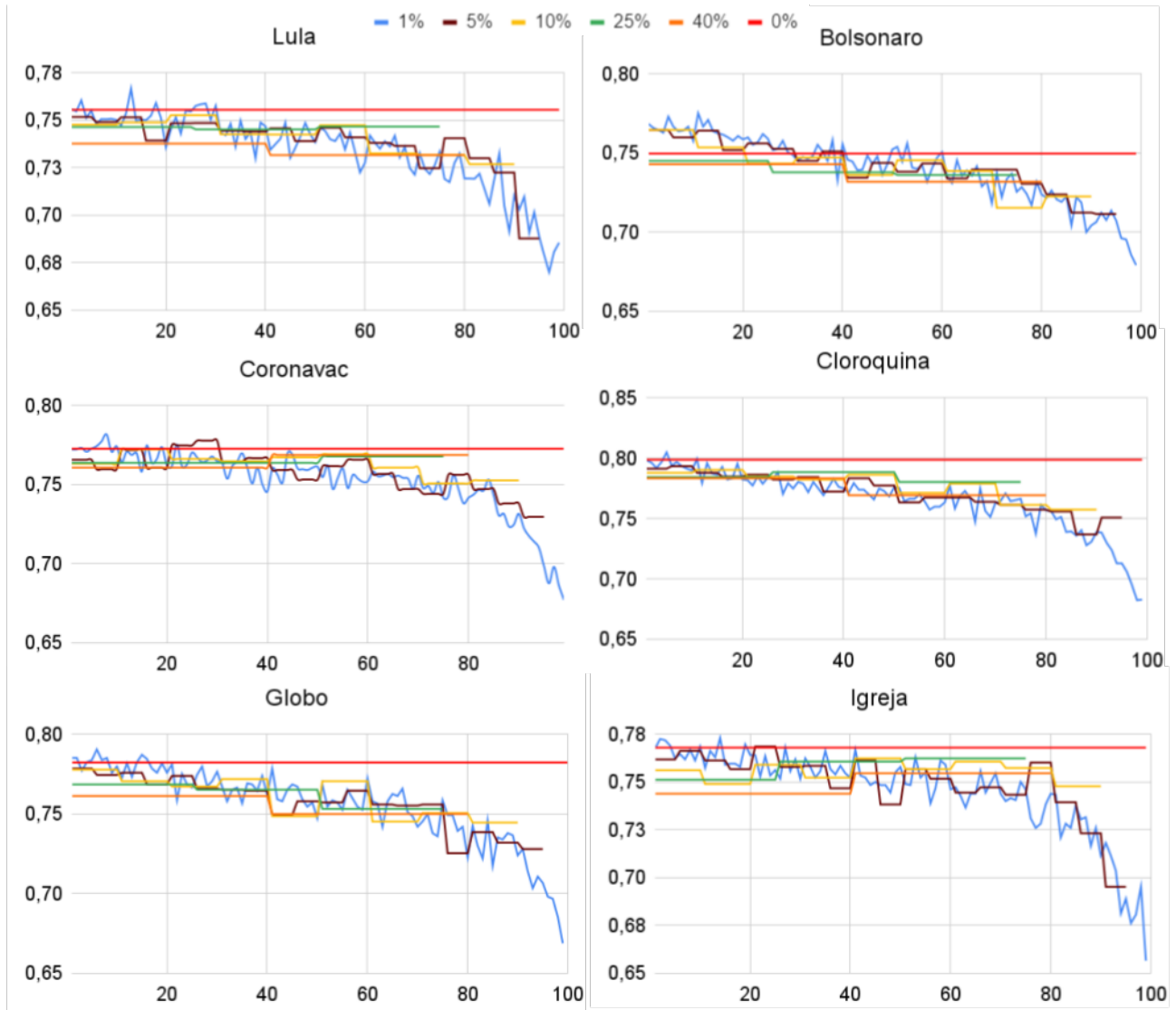


Figura 2: Expansão usando RegLog.word (horizontal: número de iterações; vertical: medida F_1).

similares aos relatados na seção 4.1, utilizando-se o córpus expandido como conjunto de treinamento para a tarefa de detecção de posicionamentos. O objetivo desses experimentos foi o de comparar o resultado da predição de posicionamento com e sem dados expandidos, tomando-se por base o mesmo conjunto de teste fixo.

A Tabela 8 sumariza os resultados obtidos com uso de cada um dos três modelos para cada um dos seis alvos do córpus. Os resultados originais do modelo *BERT*, que são os mesmos já reportados na Tabela 4 da seção 4.1, são aqui reproduzidos apenas para comparação com os resultados da expansão.

Os resultados apresentados na Tabela 8 sugerem que a inclusão de instâncias rotuladas de forma semiautomática não prejudicou o desempenho dos classificadores. Na verdade, algumas classes mostram inclusive um aumento discreto em medida F_1 médio, embora os modelos origi-

nais já tenham sido otimizados para o conjunto de dados atual. Este resultado é consistente com os obtidos pelo *framework* de expansão de córpus de análise de sentimentos CasSUL em Brum & das Graças Volpe Nunes (2018).

5.2.2. Avaliação Humana

Em complemento à avaliação intrínseca descrita na seção anterior, foi conduzida também uma avaliação humana dos dados rotulados de forma semiautomática utilizando-se o modelo *BERT* como referência. Para este fim, três juízes foram solicitados a rotular manualmente um conjunto de *tweets* aleatoriamente selecionados a partir de cada um dos melhores conjuntos de cada tópico usando as opções ‘contrário’, ‘favorável’ ou ‘outro’. Embora os dados expandidos a serem avaliados fossem totalmente rotulados com informação de classe binária, esta terceira opção foi ofere-

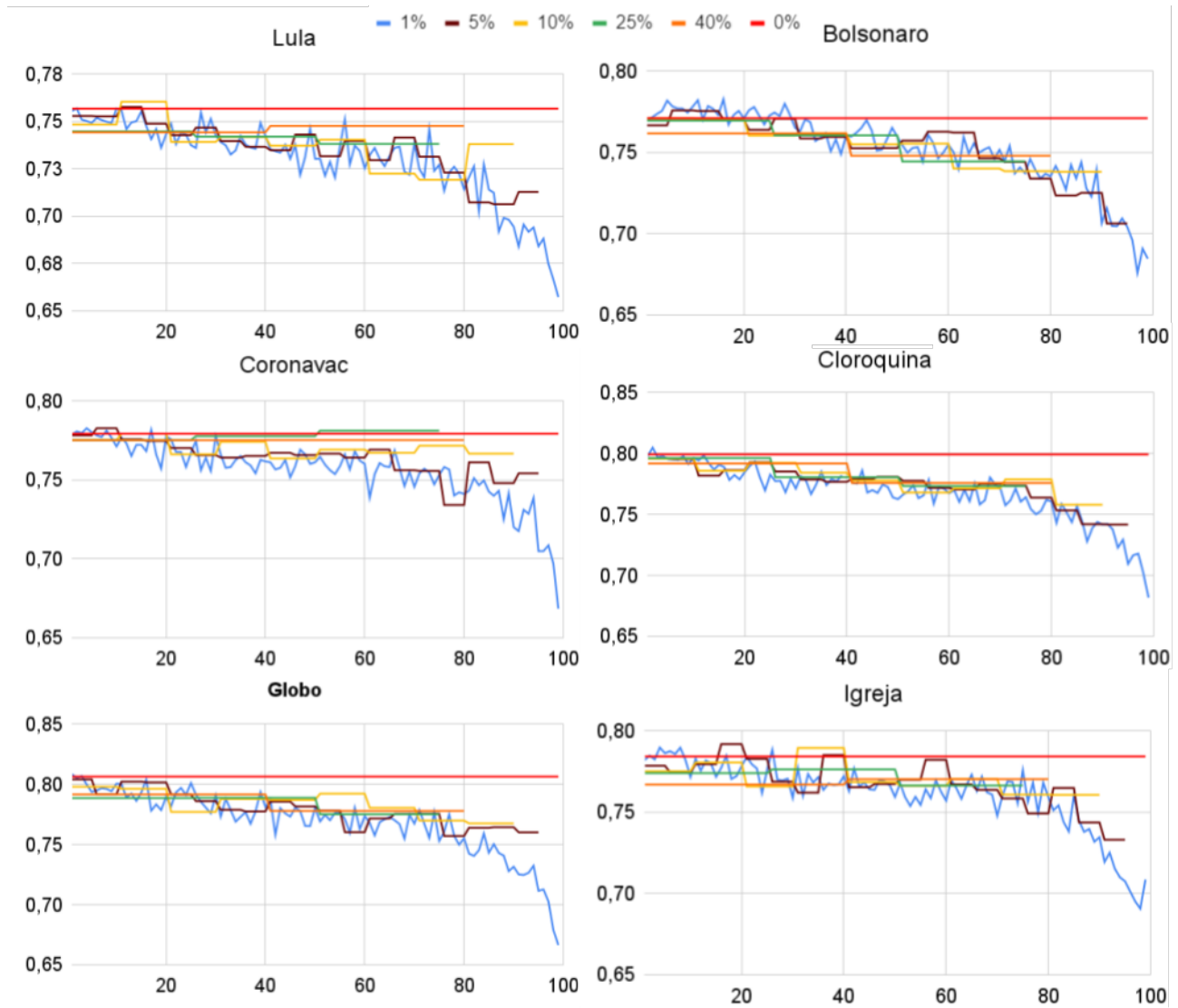


Figura 3: Expans o usando RegLog.char (horizontal: n mero de itera  es; vertical: medida F_1 .)

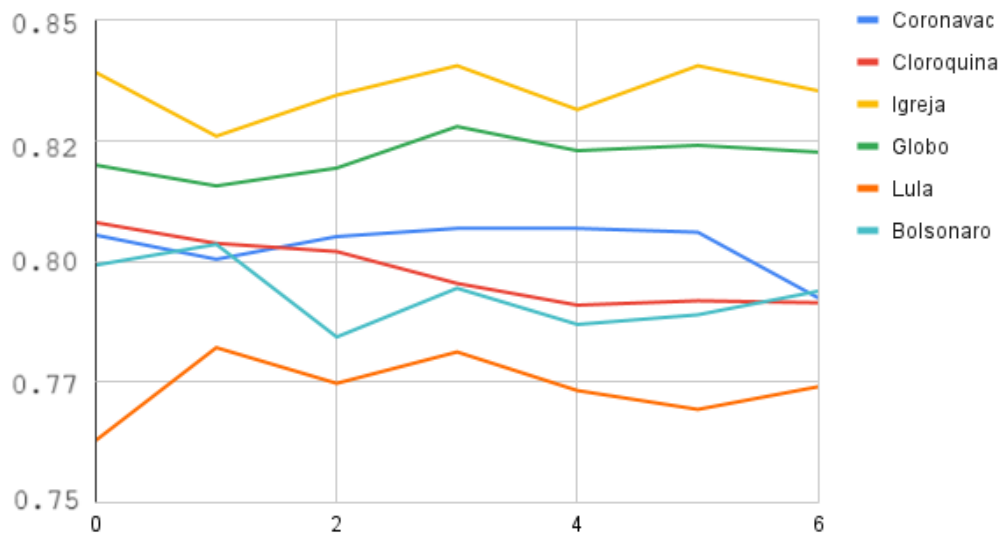


Figura 4: Expans o de dados usando BERT (horizontal: n mero de itera  es; vertical: medida F_1).

Modelo	Alvo	F_1 Original	F_1 Expandido	Novas inst.
RegLog.word	Bolsonaro	0,75	0,78	3.520
	Lula	0,76	0,77	3.575
	Coronavac	0,77	0,78	2.864
	Cloroquina	0,80	0,80	1.760
	Globo	0,78	0,79	1.794
	Igreja	0,77	0,77	4.844
RegLog.char	Bolsonaro	0,77	0,78	5.120
	Lula	0,76	0,76	5.498
	Coronavac	0,78	0,78	3.578
	Cloroquina	0,80	0,80	704
	Globo	0,81	0,81	299
	Igreja	0,78	0,79	5.187
BERT	Bolsonaro	0,80	0,80	320
	Lula	0,76	0,78	275
	Coronavac	0,81	0,81	1.074
	Cloroquina	0,80	0,80	352
	Globo	0,81	0,83	897
	Igreja	0,84	0,85	1.730

Tabela 8: Predição de posicionamentos com base em dados originais e expandidos e quantidade de novas instâncias adicionadas pelo método de expansão.

cida aos anotadores como forma de sinalizar casos ambíguos ou que não se aplicavam às categorias a favor/contra, acarretando assim uma penalidade obrigatória para o classificador binário que não conta com esta possibilidade.

A Tabela 9 apresenta o percentual de concordância entre os anotadores e o modelo de expansão, considerando-se os dados rotulados como ‘outros’ como discordantes.

Alvo	% Concordância	% Outros
Bolsonaro	90,0	5,0
Lula	86,0	9,0
Coronavac	69,0	2,0
Cloroquina	97,0	0,0
Globo	91,0	9,0
Igreja	81,0	12,0

Tabela 9: Concordância entre anotador humano e modelo de expansão na avaliação manual e número de instâncias rotuladas como ‘outros’.

Os resultados da Tabela 9 demonstram um bom grau de concordância para todos os tópicos, exceto ‘Coronavac’, em que observou-se ambiguidade significativa em alguns dos textos da amostra selecionada. Um exemplo deste tipo de ambiguidade é observado a seguir.

‘Que boa notícia! Espero mesmo que todos sejamos vacinados ainda em janeiro! Porém, uma curiosidade: alguém sabe se ou-

tro país também vai aplicar a Coronavac na população?’

Neste caso, observa-se que o posicionamento pode ser interpretado tanto como sendo favorável ao uso da vacina Coronavac (que já estaria disponível no mês de janeiro), quanto desfavorável (pois levanta dúvidas sobre a sua confiabilidade ao questionar se outros países também a adotariam).

Para melhor avaliar as discrepâncias entre as rotulações humanas e as produzidas pelo método apresentado, foi realizada também uma segunda anotação independente apenas destes casos problemáticos. Os resultados de concordância entre o avaliador humano original e este segundo juiz, representado pela primeira autora do presente trabalho, são apresentados na Tabela 10.

Alvo	% Concordância
Bolsonaro	50,0
Lula	85,7
Coronavac	29,0
Cloroquina	0,0
Globo	0,0
Igreja	36,8

Tabela 10: Concordância entre avaliador original e a autora da pesquisa para os casos de rotulação discrepantes.

É possível observar neste caso um grau razoável de concordância para os tópicos ‘Bolsonaro’ e ‘Lula’, o que talvez seja explicado pela maior polarização de opiniões a respeito destes dois tópicos, o que pode tornar estes posicionamentos mais facilmente detectáveis. Já a baixa concordância observada para os outros tópicos, embora influenciada pela inclusão de uma classe ‘outros’ não disponível ao classificador binário, sugere que a tarefa de classificação de posicionamento pode exigir grande esforço de interpretação mesmo para avaliadores humanos.

6. Considerações finais

Este artigo apresentou um método de expansão de corpus de posicionamentos em português brasileiro baseado em um *framework* existente para a tarefa correlata de análise de sentimentos (Brum & das Graças Volpe Nunes, 2018). O corpus expandido foi submetido a avaliação intrínseca e humana, com resultados indicando que a adição de instâncias rotuladas automaticamente ao corpus original não diminui a precisão da classificação, e que as instâncias adicionadas estão em sua maioria corretas.

O presente trabalho deixa uma série de oportunidades de melhoria. Dentre estas, observamos que a anotação humana tem sido utilizada apenas como ponto de partida do procedimento de expansão. No entanto, poderia ser útil, ou talvez até necessário, incluir também uma etapa de avaliação humana no ciclo de detecção-seleção, o que ajudaria a evitar a inclusão de ruído no classificador subsequente. Além disso, seria interessante comparar o presente método com alternativas recentes de anotação semiautomática de posicionamentos, fazendo uso de grandes modelos de língua, a exemplo de Zhang et al. (2023a). Iniciativas de pesquisa deste tipo são deixadas como possíveis trabalhos futuros.

Agradecimentos

Esse trabalho conta com apoio FAPESP #2021/08213-0.

Uma versão resumida deste artigo foi apresentada em Pereira & Paraboni (2024), à qual foi acrescentada uma revisão mais completa da área de detecção de posicionamentos a partir de texto (seção 2), novos experimentos de detecção de posicionamentos com uma visão mais detalhada da importância das características de aprendizado (seção 4), e uma descrição mais aprofundada dos experimentos de expansão e seus re-

sultados, incluindo novos modelos de *baseline*, a apresentação dos resultados em forma gráfica e mais considerações sobre a tarefa de anotação humana (seção 5).

Referências

- Aldayel, Abeer & Walid Magdy. 2021. Stance detection on social media: State of the art and trends. *Information Processing & Management* 58(4). doi 10.1016/j.ipm.2021.102597
- Allaway, Emily & Kathleen R. McKeown. 2020. Zero-shot stance detection: A dataset and model using generalized topic representations. Em *Empirical Methods in Natural Language Processing (EMNLP)*, 8913–8931. doi 10.18653/v1/2020.emnlp-main.717
- Ayyub, Kashif, Saqib Iqbal, Muhammad Wasif Nisar, Saima Gulzar Ahmad & Ehsan Ullah Munir. 2021. Stance detection using diverse feature sets based on machine learning techniques. *Journal of Intelligent & Fuzzy Systems* 40(5). 9721–9740. doi 10.3233/JIFS-202269
- Breiman, Leo. 2001. Random forests. *Machine Learning* 45(1). 5–32. doi 10.1023/A:1010933404324
- Brum, Henrico Bertini & Maria das Graças Volpe Nunes. 2018. Semi-supervised sentiment annotation of large corpora. Em *Conference on Computational Processing of Portuguese PROPOR*, 385–395. doi 10.1007/978-3-319-99722-3_39
- Bugueño, Margarita & Marcelo Mendoza. 2019. Applying self-attention for stance classification. Em *Iberoamerican Congress on Pattern Recognition (CIARP)*, 51–61. doi 10.1007/978-3-030-33904-3_5
- Carvalho, Cecília Silvestre, Vlória C Pinheiro & Lívio Freire. 2020. A deep learning model of common sense knowledge for augmenting natural language processing tasks in Portuguese language. Em *Computational Processing of the Portuguese Language*, 303–312. doi 10.1007/978-3-030-41505-1_29
- Cavalheiro, Laís Carraro Leme, Matheus Camasmie Pavan & Ivandré Paraboni. 2023. Stance prediction from multimodal social media data. Em *Recent Advances in Natural Language Processing (RANLP)*, 242–248. doi 10.26615/978-954-452-092-2_027
- Chen, Ninghan, Xihui Chen & Jun Pang. 2022. A multilingual dataset of COVID-19 vaccination attitudes on Twitter. *Data in Brief* 44. 108503. doi 10.1016/j.dib.2022.108503

- Chunling, Wang, Zhang Yijia, Yu Xingyu, Liu Guantong, Chen Fei & Lin Hongfei. 2023. Adversarial network with external knowledge for zero-shot stance detection. Em *22nd Chinese National Conference on Computational Linguistics*, 824–835. doi 10.1007/978-981-99-6207-5_26
- Conforti, Costanza, Jakob Berndt, Marco Basaldella, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd & Nigel Collier. 2021. Adversarial training for news stance detection: Leveraging signals from a multi-genre corpus. Em *EACL Hackashop on News Media Content Analysis and Automated Report Generation*, 1–7. 
- Conforti, Costanza, Jakob Berndt, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd & Nigel Collier. 2020. Will-They-Won't-They: A very large dataset for stance detection on Twitter. Em *58th Meeting of the Association for Computational Linguistics (ACL)*, 1715–1724. doi 10.18653/v1/2020.acl-main.157
- da Costa, Pablo Botton, Matheus Camasmie Pavan, Wesley Ramos dos Santos, Samuel Caetano da Silva & Ivandré Paraboni. 2023. BERTabaporu: assessing a genre-specific language model for Portuguese NLP. Em *Recent Advances in Natural Language Processing (RANLP)*, 217–223. doi 10.26615/978-954-452-092-2_024
- D'Andrea, Eleonora, Pietro Ducange, Alessio Bechini, Alessandro Renda & Francesco Marcelloni. 2019. Monitoring the public opinion about the vaccination topic from tweets analysis. *Expert Systems with Applications* 116. 209–226. doi 10.1016/j.eswa.2018.09.009
- Darwish, Kareem, Walid Magdy & Tahar Zanoua. 2017. Improved stance prediction in a user similarity feature space. Em *Advances in Social Network Analysis and Mining (ASONAM)*, 145–148. doi 10.1145/3110025.3110112
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee & Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. Em *North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, 4171–4186. doi 10.18653/v1/N19-1423
- Fang, Wei, Moin Nadeem, Mitra Mohtarami & James Glass. 2019. Neural multi-task learning for stance prediction. Em *2nd Workshop on Fact Extraction and VERification (FEVER)*, 13–19. doi 10.18653/v1/D19-6603
- Flores, Arthur Marçal, Matheus Camasmie Pavan & Ivandré Paraboni. 2022. User profiling and satisfaction inference in public information access services. *Journal of Intelligent Information Systems* 58(1). 67–89. doi 10.1007/s10844-021-00661-w
- Geiss, Henri-Jacques, Flora Sakketou & Lucie Flek. 2022. OK boomer: Probing the socio-demographic divide in echo chambers. Em *10th International Workshop on Natural Language Processing for Social Media*, 83–105. doi 10.18653/v1/2022.socialnlp-1.8
- Glandt, Kyle, Sarthak Khanal, Yingjie Li, Doina Caragea & Cornelia Caragea. 2021. Stance detection in COVID-19 Tweets. Em *59th Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 1596–1611. doi 10.18653/v1/2021.acl-long.127
- Hanawa, Kazuaki, Akira Sasaki, Naoaki Okazaki & Kentaro Inui. 2019. Stance detection attending external knowledge from Wikipedia. *Journal of Information Processing* 27. 499–506. doi 10.2197/ipsjjip.27.499
- Hassan, Fuad Mire & Mark Lee. 2019. Imbalanced stance detection by combining neural and external features. Em *Statistical Language and Speech Processing (SLSP)*, 273–285. doi 10.1007/978-3-030-31372-2_23
- Jaziriyani, Mohammad Mehdi, Ahmad Akbari & Hamed Karbasi. 2021. ExaASC: A general target-based stance detection corpus in Arabic language. Em *11th International Conference on Computer Engineering and Knowledge (ICCKE)*, 424–429. doi 10.1109/ICCKE54056.2021.9721486
- Kovačević, Aleksandar, Olivera Grljević, Zita Bošnjak & Gordana Svilengaćin. 2020. The linguistic construction of sentiment expressions in student opinionated content: A corpus-based study. *Poznan Studies in Contemporary Linguistics* 56(2). 207–249. doi 10.1515/psic1-2020-0006
- Küçük, Dilek & Fazli Can. 2020. Stance detection: A survey. *ACM Computing Surveys* 53(1). 1–37. doi 10.1145/3369026
- Lai, Mirko, Alessandra Teresa Cignarella, Delia Irazú Hernandez Farias, Cristina Bosco, Viviana Patti & Paolo Rosso. 2020. Multilingual stance detection in social media political debates. *Computer Speech and Language* 63. 101075. doi 10.1016/j.cs1.2020.101075

- Li, Ang, Bin Liang, Jingqian Zhao, Bowen Zhang, Min Yang & Ruifeng Xu. 2023. Stance detection on social media with background knowledge. Em *Empirical Methods in Natural Language Processing (EMNLP)*, 15703–15717. doi 10.18653/v1/2023.emnlp-main.972
- Li, Yingjie & Cornelia Caragea. 2019. Multi-task stance detection with sentiment and stance lexicons. Em *Empirical Methods in Natural Language Processing and International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 6299–6305. doi 10.18653/v1/D19-1657
- Li, Yingjie, Tiberiu Sosea, Aditya Sawant, Ajith Jayaraman Nair, Diana Inkpen & Cornelia Caragea. 2021. P-stance: A large dataset for stance detection in political domain. Em *Findings of the Association for Computational Linguistics*, 2355–2365. doi 10.18653/v1/2021.findings-acl.208
- Magdy, Walid, Kareem Darwish, Norah Abokhodair, Afshin Rahimi & Timothy Baldwin. 2016. #ISISisNotIslam or #DeportAllMuslims? predicting unspoken views. Em *8th Conference on Web Science*, 95–106. doi 10.1145/2908131.2908150
- Mayfield, Elijah & Alan W Black. 2019. Stance classification, outcome prediction, and impact assessment: NLP tasks for studying group decision-making. Em *3rd Workshop on Natural Language Processing and Computational Social Science*, 65–77. doi 10.18653/v1/W19-2108
- Mohammad, Saif, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu & Colin Cherry. 2016. SemEval-2016 task 6: Detecting stance in tweets. Em *10th International Workshop on Semantic Evaluation (SemEval-2016)*, 31–41. San Diego, California: Assoc. for Computational Linguistics. doi 10.18653/v1/S16-1003
- Mutlu, Ece C., Toktam Oghaz, Jasser Jasser, Ege Tutunculer, Amirarsalan Rajabi, Aida Tayebi, Ozlem Ozmen & Ivan Garibay. 2020. A stance data set on polarized conversations on twitter about the efficacy of hydroxychloroquine as a treatment for COVID-19. *Data in brief* 33(106401). doi 10.1016/j.dib.2020.106401
- Ng, Lynnette Hui Xian & Kathleen M. Carley. 2022. Is my stance the same as your stance? a cross validation study of stance detection datasets. *Information Processing & Management* 59(6). 103070. doi 10.1016/j.ipm.2022.103070
- Pavan, Matheus Camasmie & Ivandr  Paraboni. 2022. Cross-target stance classification as domain adaptation. Em *21st Mexican International Conference on Artificial Intelligence (MICAI)*, 15–25. doi 10.1007/978-3-031-19493-1_2
- Pavan, Matheus Camasmie, Vitor Garcia dos Santos, Alex Gwo Jen Lan, Jo  o Trevisan Martins, Wesley Ramos dos Santos, Caio Deutsch, Pablo Botton da Costa, Fernando Chiu Hsieh & Ivandr  Paraboni. 2023. Morality classification in natural language text. *IEEE transactions on Affective Computing* 14(1). 857–863. doi 10.1109/TAFFC.2020.3034050
- Pavan, Matheus Camasmie, Wesley Ramos dos Santos & Ivandr  Paraboni. 2020. Twitter moral stance classification using long short-term memory networks. Em *Brazilian Conference on Intelligent Systems (BRACIS)*, 636–647. doi 10.1007/978-3-030-61377-8_45
- Pereira, Camila & Ivandr  Paraboni. 2024. Semi-automatic corpus expansion: the case of stance prediction. Em *16th International Conference on Computational Processing of Portuguese (PROPOR)*, 515–520.  
- Pereira, Camila, Matheus Pavan, Sungwon Yoon, Ricelli Ramos, Pablo Costa, La s Cavalcheiro & Ivandr  Paraboni. 2023. UstanceBR: a multimodal language resource for stance prediction. ArXiv [cs.CL]. doi 10.48550/arXiv.2312.06374
- Popat, Kashyap, Subhabrata Mukherjee, Andrew Yates & Gerhard Weikum. 2019. STANCY: stance classification based on consistency cues. ArXiv [cs.CL]. doi 10.48550/arXiv.1910.06048
- Samih, Younes & Kareem Darwish. 2021. A few topical Tweets are enough for effective user stance detection. Em *16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2637–2646. doi 10.18653/v1/2021.eacl-main.227
- Sane, Sushmitha Reddy, Suraj Tripathi, Koushik Reddy Sane & Radhika Mamidi. 2019. Stance detection in code-mixed Hindi-English social media data using multi-task learning. Em *10th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 1–5. doi 10.18653/v1/W19-1301
- dos Santos, Wesley Ramos & Ivandr  Paraboni. 2019. Moral stance recognition and

- polarity classification from Twitter and elicited text. Em *Recent Advances in Natural Language Processing (RANLP)*, 1069–1075. doi 10.26615/978-954-452-056-4_123
- Schaefer, Robin & Manfred Stede. 2019. Improving implicit stance classification in tweets using word and sentence embeddings. Em *42nd German Conference on Artificial Intelligence*, 299–307. doi 10.1007/978-3-030-30179-8_26
- Siddiqua, Umme Aymun, Abu Nowshed Chy & Masaki Aono. 2019. Tweet stance detection using an attention based neural ensemble model. Em *Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 1868–1873. doi 10.18653/v1/N19-1185
- da Silva, Samuel Caetano, Thiago Castro Ferreira, Ricelli Moreira Silva Ramos & Ivandré Paraboni. 2020. Data driven and psycholinguistics motivated approaches to hate speech detection. *Computación y Sistemas* 24(3). 1179–1188. doi 10.13053/CyS-24-3-3478
- Sobhani, Parinaz, Diana Inkpen & Xiaodan Zhu. 2017. A dataset for multi-target stance detection. Em *15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 551–557. [↗](#)
- de Sousa, André & Karin Becker. 2023. SSSD: Leveraging pre-trained models and semantic search for semi-supervised stance detection. Em *14th International Conference on Recent Advances in Natural Language Processing (RANLP)*, 264–273. [↗](#)
- Su, Zhizhong, Yaoyi Xi, Rong Cao, Huifeng Tang & Hangyu Pan. 2021. A stance detection approach based on generalized autoregressive pretrained language model in Chinese microblogs. Em *13th 13th International Conference on Machine Learning and Computing*, 232–238. doi 10.1145/3457682.3457717
- Taulé, Mariona, Maria Antònia Martí, Francisco Manuel Rangel Pardo, Paolo Rosso, Cristina Bosco & Viviana Patti. 2017. Overview of the task on stance and gender detection in Tweets on Catalan Independence at. Em *2nd Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval)*, 157–177
- Tun, Yin Min & Phyu Hninn Myint. 2019. A two-phase approach for stance classification in Twitter using name entity recognition and term frequency feature. Em *18th 18th International Conference on Computer and Information Science (ICIS)*, 77–81. doi 10.1109/ICIS46139.2019.8940282
- Wen, Haoyang & Alexander Hauptmann. 2023. Zero-shot and few-shot stance detection on varied topics via conditional generation. Em *61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 1491–1499. doi 10.18653/v1/2023.acl-short.127
- Xu, Chang, Cecile Paris, Surya Nepal & Ross Sparks. 2019a. Recognising agreement and disagreement between stances with reason comparing networks. *arXiv preprint arXiv:1906.01392*
- Xu, Zhenhui, Qiang Li, Wei Chen, Yingbao Cui, Zhen Qiu & Tengjiao Wang. 2019b. Opinion-aware knowledge embedding for stance detection. Em *Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data*, 337–348. Springer
- Zhang, Bowen, Daijun Ding, Zhichao Huang, Ang Li, Yangyang Li, Baoquan Zhang & Hu Huang. 2024. Knowledge-augmented interpretable network for zero-shot stance detection on social media. *IEEE Transactions on Computational Social Systems* 1–12. doi 10.1109/TCSS.2024.3388723
- Zhang, Bowen, Daijun Ding & Liwen Jing. 2023a. How would stance detection techniques evolve after the launch of ChatGPT? *ArXiv [cs.CL]*. doi 10.48550/arXiv.2212.14548
- Zhang, Bowen, Daijun Ding, Guangning Xu, Jinjin Guo, Zhichao Huang & Xu Huang. 2023b. Twitter stance detection via neural production systems. Em *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. doi 10.1109/ICASSP49357.2023.10094597
- Zhang, Lei & Bing Liu v Wang. 2018. Deep learning for sentiment analysis: A survey. *WIRES Data Mining and Knowledge Discovery* 8(4). e1253. doi 10.1002/widm.1253
- Zhao, Xuechen, Jiaying Zou, Zhong Zhang, Feng Xie, Bin Zhou & Lei Tian. 2022. Zero-shot stance detection based on cross-domain feature enhancement by contrastive learning. *ArXiv [cs.CL]*. doi 10.48550/arXiv.2210.03380
- Zhou, Shengping, Junjie Lin, Lianzhi Tan & Xin Liu. 2019. Condensed convolution neural network by attention over self-attention for stance detection in

- twitter. Em *International Joint Conference on Neural Networks (IJCNN)*, 1–8.
 [10.1109/IJCNN.2019.8851965](https://doi.org/10.1109/IJCNN.2019.8851965)
- Zhu, Xiaojin Jerry. 2005. Semi-supervised learning literature survey. Relatório técnico. University of Wisconsin-Madison Department of Computer Sciences. [↗](#)
- Zhu, Zhengyuan, Kevin Meng, Josue Caraballo, Israa Jaradat, Xiao Shi, Zeyu Zhang, Farahnaz Akrami, Haojin Liao, Fatma Arslan, Damian Jimenez et al. 2021. A dashboard for mitigating the COVID-19 misinfodemic. Em *16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 99–105.
 [10.18653/v1/2021.eacl-demos.12](https://doi.org/10.18653/v1/2021.eacl-demos.12)