

# UM ESTUDO DE PISTAS LINGÜÍSTICAS PARA DETECTAR NOTÍCIAS FALSAS ESCRITAS EM LÍNGUA PORTUGUESA

RODRIGO RODRIGUES<sup>1</sup>; LARISSA A. DE FREITAS<sup>2</sup>

<sup>1</sup>Universidade Federal de Pelotas – [rfrdrigues@inf.ufpel.edu.br](mailto:rfrdrigues@inf.ufpel.edu.br)

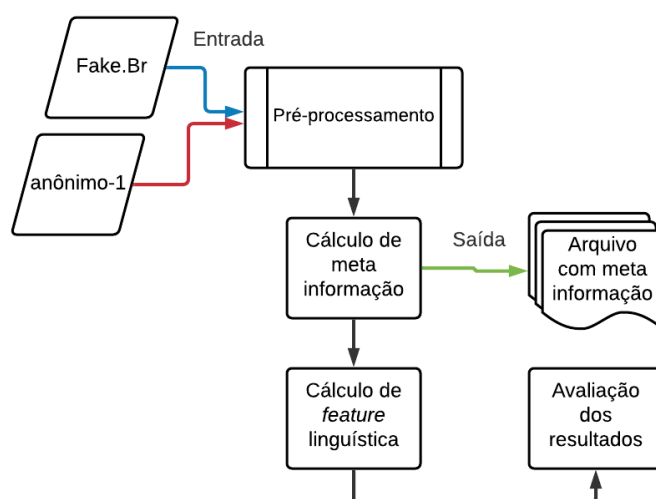
<sup>2</sup>Universidade Federal de Pelotas – [larissa@inf.ufpel.edu.br](mailto:larissa@inf.ufpel.edu.br)

## 1. INTRODUÇÃO

Hoje em dia, o volume de conteúdo gerado a cada minuto é imenso, principalmente nas redes sociais, que facilitam o compartilhamento rápido de informações. Com isso, surge um grande problema, a verificação da veracidade das informações compartilhadas. Seres humanos, não são muito bons em detectar mentiras, em parte por causa do viés natural da verdade (ZHOU *et al.*, 2008). Neste contexto, destaca-se uma área do Processamento de Língua Natural (PLN), denominada detecção de notícias falsas, a qual pode ser realizada através de abordagens que fazem uso de *features* linguísticas (MONTEIRO *et al.*, 2018; ZHOU *et al.*, 2008; ZHOU *et al.*, 2004). Pistas linguísticas são comportamentos realizados por um transmissor que sugerem a probabilidade de que esteja tentando enganar (ZHOU *et al.*, 2008). Inspirado em iniciativas anteriores, este artigo tem como objetivo apresentar uma ferramenta que extrai *features* linguísticas de corpora sobre notícias falsas.

## 2. METODOLOGIA

Neste trabalho, construímos uma ferramenta que extrai *features* linguísticas de corpora sobre notícias falsas. Em nosso experimento usamos dois *datasets*, Fake.Br e anônimo-1. A Figura 1 apresenta o fluxograma da ferramenta LC-tool.<sup>1</sup>



**Figura 1** - Fluxograma da Ferramenta LC-tool. Fonte: Própria.

No pré-processamento removemos os caracteres de controle e espaços extras, terminando com uma lista de tokens limpos, útil para a etapa de cálculo de

<sup>1</sup> <https://github.com/rferreirarodrigues/LC-tool>

meta informação. Na etapa de cálculo de meta informação, todas as informações necessárias para o cálculo das *features* são calculadas (i.e tokens, tipos, links, *sentenças*, verbos, etc). Para cada notícia, um arquivo de meta informação é gerado.

## 2.1 DATASETS

Os *datasets* usados neste trabalho foram o Fake.Br (MONTEIRO *et al.*, 2018), criado pelo grupo de pesquisa NILC, composto por 3600 notícias verdadeiras e 3600 notícias falsas; e o anônimo-1<sup>2</sup>, criado pelo grupo de pesquisa MIDIARS, composto por 8 notícias classificadas como desinformação, 19 como desmentido e 23 como outro. Ambos corpora são de domínio público.

## 2.2 FEATURES LINGUÍSTICAS

Com base na conclusão de outro trabalho (ZHOU *et al.*, 2008), a Tabela 1 apresenta um conjunto de *features* linguísticas que foram usados para encontrar pistas úteis para detectar notícias falsas.

**Tabela 1.** *Features* linguísticas.

<b>Feature linguística e sua descrição</b>	<b>Descobertas</b>
<b>quantidade de informações:</b> número de tokens, palavras, adjetivos, advérbios e sentenças	mensagens enganosas usam menos informações
<b>complexidade da linguagem:</b> número médio de sentenças, número médio de palavras, pausalidade	mensagens enganosas são mais complexas
<b>não-imediatismo:</b> expressão para se dissociar o conteúdo de sua mensagem	mensagens enganosas apresentam maior grau de não-imediatismo
<b>expressividade:</b> emotividade	mensagens enganosas possuem maior expressividade
<b>diversidade:</b> diversidade lexical, porcentagem de palavras únicas	mensagens enganosas possuem menor diversidade
<b>informalidade:</b> média de palavras com erros ortográficos	mensagens enganosas possuem maior grau de informalidade

Fonte: Lina Zhou, Dongsong Zhang Communications of the ACM, September 2008, Vol. 51 No. 9, Pages 119-122

## 3. RESULTADOS E DISCUSSÃO

A Tabela 2 apresenta os resultados obtidos para os dois *datasets*, Fake.Br e anônimo-1.

<sup>2</sup> <https://wp.ufpel.edu.br/midiars/datasets/>

**Tabela 2.** Resultados das *features* linguísticas aplicadas nos *datasets Fake.Br* e *anônimo-1*.

<i>Features</i>	<i>Fake</i>	<i>True</i>	<i>Desme ntido</i>	<i>Desinfo rmação</i>	<i>Outro</i>
Número de verbos modais	<b>24153</b>	128965	9	<b>7</b>	13
Número médio de tokens	215.47	1268.09	10.38	4.97	11.90
Número médio de tipos <sup>3</sup> (sem pontuação e números)	<b>125.16</b>	501.20	9.65	<b>4.58</b>	11.18
Número médio de palavras (em caracteres)	<b>5.07</b>	5.09	5.59	<b>4.92</b>	5.19
Razão tipo-token	0.58	0.39	0.92	0.92	0.93
Número médio de sentenças	<b>12.19</b>	58.61	1.10	<b>0.55</b>	1.25
Número médio de frases (em palavras)	15.09	18.66	8.05	7.86	8.28
Número médio de verbos (norm. pelo número médio de tokens)	14.55	13.42	13.73	19.60	13.45
Número médio de substantivos (norm. pelo número médio de tokens)	28.93	29.04	34.22	26.63	32.77
Número médio de adjetivos (norm. pelo número médio de tokens)	4.22	4.37	3.13	2.01	5.25
Número médio de advérbios (norm. pelo número médio de tokens)	3.74	3.71	3.13	3.02	3.15
Número médio de pronomes (norm. pelo número médio de tokens)	12.56	12.58	12.05	18.59	10.08
Número médio de stopwords (norm. pelo número médio de tokens)	40.28	42.39	36.14	41.71	34.87
Média de palavras com erros ortográficos	<b>0.44</b>	0.38	1.98	2.89	4.35
Pausalidade média por texto	2.59	2.98	1.39	1.18	1.24
Emotividade média por texto	0.18	0.19	0.13	0.11	0.18
Não imediatismo média por texto	0.53	3.25	0.00	0.03	0.00
Diversidade média por texto	0.68	0.46	1.09	1.06	1.08

<sup>3</sup> Palavras exclusivas contidas em uma notícia.

Os enganadores apresentam ansiedade cognitiva quanto à possibilidade de serem detectados (ZHOU *et al.*, 2008). Eles podem inadvertidamente adotar um grau mais alto de não imediatismo e informalidade em suas mensagens do que os contadores da verdade. A informalidade observada na Tabela 2 está em maior grau em notícias falsas, 0.44 no Fake.Br.

É importante ressaltar que o *dataset* anônimo-1 é desbalanceado, a maioria das notícias foram pré-classificadas com o rótulo outro, impactando no resultado de algumas *features*, como: quantidade de informações, informalidade e diversidade.

Ainda na Tabela 2, observou-se que as notícias falsas apresentam menos palavras, sentenças, verbos modais, as quais são pistas linguísticas relatadas em outros estudos (ZHOU *et al.*, 2008, ZHOU *et al.*, 2008., ZHOU *et al.*, 2004., ZHOU *et al.*, 2003) que sugerem conteúdo falso.

#### 4. CONCLUSÕES

Este trabalho apresentou a ferramenta LC-Tool que extrai *features* linguísticas de corpora sobre notícias falsas. As *features* linguísticas implementadas por esta ferramenta foram propostas por Lina Zhou e Dongsong Zhang (ZHOU *et al.*, 2008). Em nossos experimentos, utilizamos dois corpora em língua portuguesa sobre notícias falsas, Fake.Br e anônimo-1. Em nossos experimentos, observamos que as notícias falsas usam menos informações: tokens, verbos modais, sentenças, substantivos e pronomes. Notícias enganosas têm menos diversidade lexical e notícias enganosas têm maior informalidade. É importante ressaltar que os corpora utilizados não foram truncados, as notícias têm tamanhos diferentes, o que é um fator que pode ter enviesado algumas *features*. Como trabalhos futuros pretendemos estender o conjunto de *features* linguísticas e usar diferentes técnicas de aprendizado de máquina para classificar as notícias como verdadeiras ou falsas.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

Monteiro R.A., Santos R.L.S., Pardo T.A.S., de Almeida T.A., Ruiz E.E.S., Vale O.A. (2018) Contributions to the Study of Fake News in Portuguese: New Corpus and Automatic Detection Results. In: 13th International Conference on the Computational Processing of Portuguese, 324–334.

Zhou, L., Twitchell, D.P., Qin, T., Burgoon, J.K., Nunamaker, J.F. (2003) An exploratory study into deception detection in text-based computer-mediated communication. In: 36th Annual Hawaii International Conference on System Sciences, 1-10

Zhou, L., Zhang, D. (2008) Following linguistic footprints: Automatic deception detection in online communication. Communications of the ACM - Enterprise Information Integration: and other tools for merging data 51(9), 119–122.

Zhou, L., Burgoon, J., Twitchell, D., Qin, T., Nunamaker Jr., J. (2004) A comparison of classification methods for predicting deception in computer-mediated communication. Journal of Management Information Systems 20(4), 139–165.