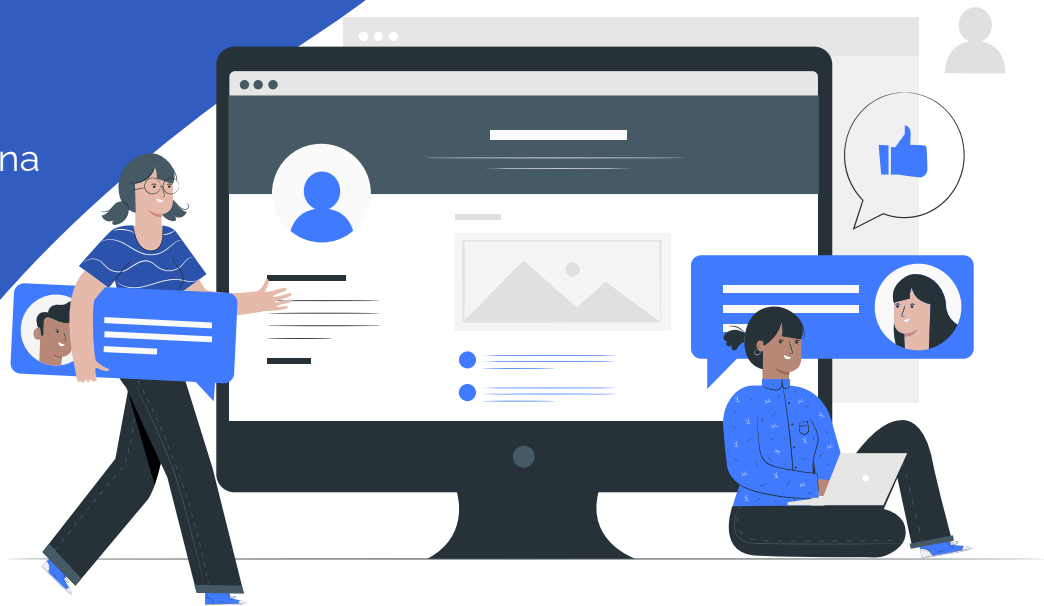


Detecção de *Fake News* em Português Utilizando Inferência de Linguagem Natural

Larissa Mayumi Barela Hondo
RA: 191026123

Orientador: Prof. Assoc. Aparecido Nilceu Marana



Introdução

Problema

- Reconhecer a veracidade das notícias tornou-se um desafio nas mídias sociais *online*;
- As *fake news* podem causar danos irreparáveis a indivíduos, organizações e governos;
- Não há muitos trabalhos utilizando a ILN aplicada à apuração de *fake news* como uma alternativa à análise das notícias por seu conteúdo e local onde foram publicadas, especialmente na língua portuguesa.

Objetivo Geral

- Método de detecção automática de *fake news* na língua portuguesa utilizando ILN;
- Dados dois textos de entrada, um sendo uma notícia já confirmada como verdadeira e outro sendo uma notícia suspeita, é retornada uma relação de inferência a fim de auxiliar a identificação de *fake news* com o uso de ILN;

Justificativa

- Possui relevância dentro da área técnica e acadêmica;
- Um dos métodos de se detectar *fake news* consiste em utilizar ILN, imitando a maneira como os especialistas analisam as informações.

Fundamentação Teórica



Processamento de Linguagem Natural

- O Processamento de Linguagem Natural (PLN) consiste em uma coleção de técnicas computacionais para automatizar a análise e representação de linguagens humanas.

Inferência de Linguagem Natural

- Os humanos usam uma variedade de conhecimentos e raciocínios para auxiliar o entendimento dos significados da linguagem;
- Inferência de Linguagem Natural (ILN) é tarefa de determinar se uma “hipótese” é verdadeira (implicação), falsa (contradição) ou indeterminada (neutra), quando comparada a uma “premissa”.

Inferência de Linguagem Natural

- Exemplos da base de dados *Stanford Natural Language Inference* (SNLI):

Exemplo de hipótese verdadeira:

Premissa: Um jogo de futebol com vários homens jogando.

Hipótese: Alguns homens estão praticando um esporte.

Exemplo de hipótese falsa:

Premissa: Um homem inspeciona o uniforme de uma figura em algum país do Leste Asiático.

Hipótese: O homem está dormindo.

Exemplo de hipótese neutra:

Premissa: Um homem mais velho e um mais jovem sorrindo.

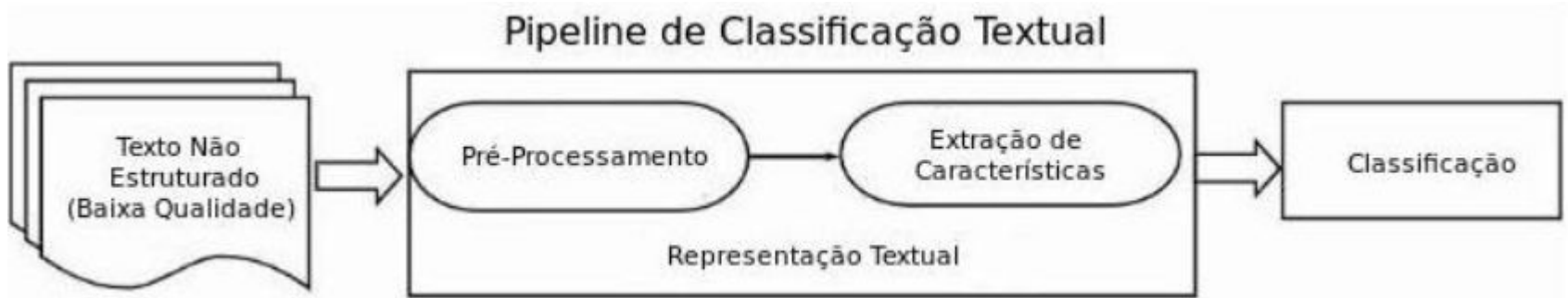
Hipótese: Dois homens estão sorrindo e rindo com os gatos brincando no chão.

Inferência de Linguagem Natural e *Fake News*

- As *fake news* são artigos fictícios fabricados deliberadamente, que podem enganar os leitores, e já existiam antes mesmo da Internet, tradição oral, em forma de rumores ou comentários acerca de outras pessoas ou de organizações rivais;
- A detecção de *fake news* por humanos baseia-se na inferência da veracidade utilizando um conjunto de notícias confiáveis;
- A notícia suspeita é considerada como uma hipótese e as notícias confiáveis como uma premissa;
- As relações de inferência entre as premissas e as hipóteses revelam as fiabilidades das notícias suspeitas.

Representação de Palavras

- O texto deve ser convertido em um formato que o computador possa entender;



Fonte: Adaptada de Naseem et al. (2020).



Ferramentas

Bibliotecas

- Googletrans;
- NLTK;
- Numpy;
- Pandas;
- Scikit-Learn;
- Hugging Face;
- TensorFlow.

Pré-processamento Textual

- Remoção de pontuações;
- Conversão de todos os caracteres para minúsculos;
- Remoção de *stop words*;
- Tokenização;
- Lematização;
- Remoção de acentos;

Extração de Características

- *Bag of Words* (BoW) transforma uma coleção de documentos de texto em vetores de características numéricas;
- Os documentos são descritos por ocorrências de palavras, ignorando completamente as informações de posição relativa das palavras no documento.

Extração de Características

- TF-IDF:

$$\text{IDF}(t) = \log \frac{n}{1 + \text{DF}(t)}$$

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t)$$

- Onde n é o número total de documentos no conjunto de documentos e $\text{DF}(t)$ é o número de documentos no conjunto de documentos que contêm o termo t , sendo que a importância de um termo é inversamente relacionada à sua frequência nos documentos

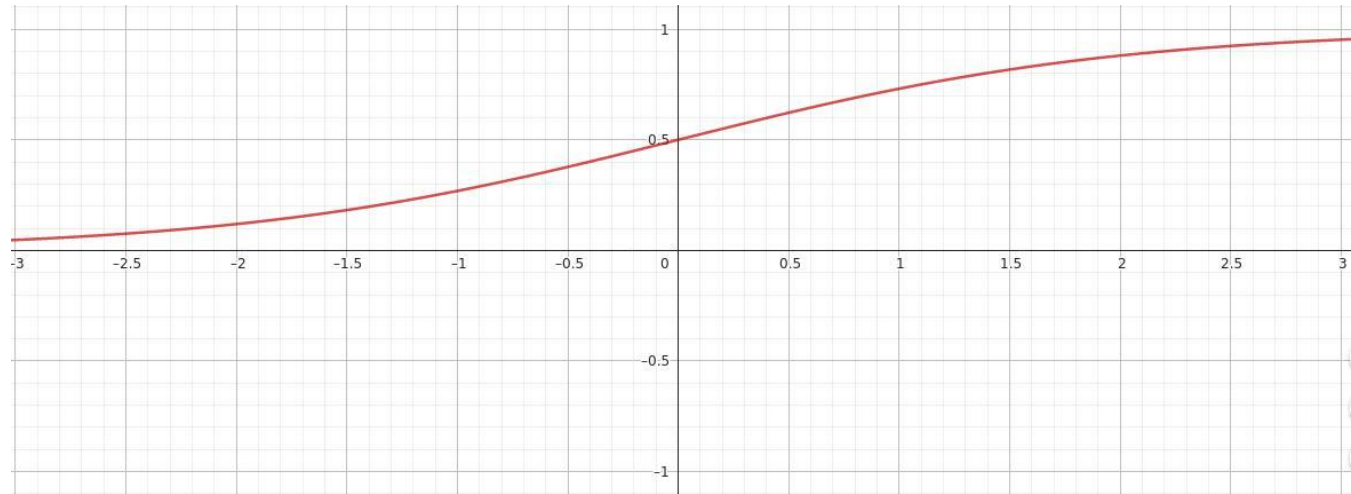
Regressão Logística (LR)

$$\text{logit}(\pi) = \ln\left(\frac{\pi}{1 - \pi}\right) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k = \alpha + \beta X$$

$$e^{\alpha + \beta X} = \frac{\pi}{1 - \pi}$$

$$e^{\alpha + \beta X} - e^{\alpha + \beta X} \pi = \pi$$

$$\pi = \frac{e^{\alpha + \beta X}}{1 + e^{\alpha + \beta X}}$$



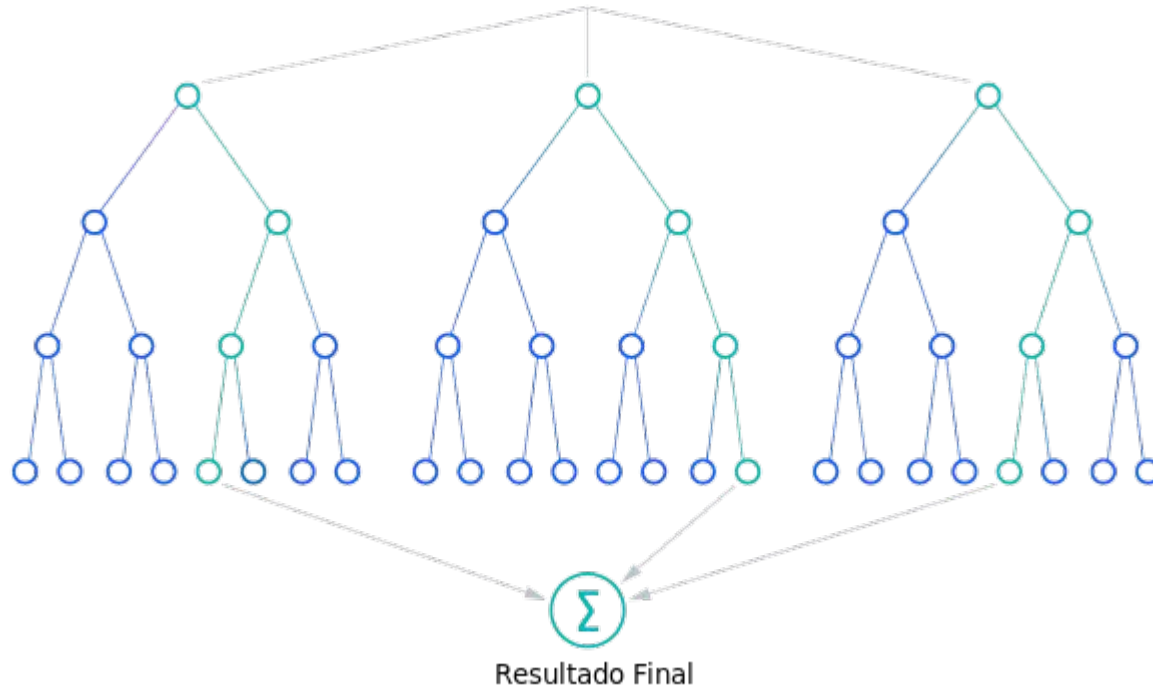
Fonte: Elaborada pela autora.

Naive Bayes (NB)

- Aplicação do teorema de Bayes de independência condicional entre cada par de características:

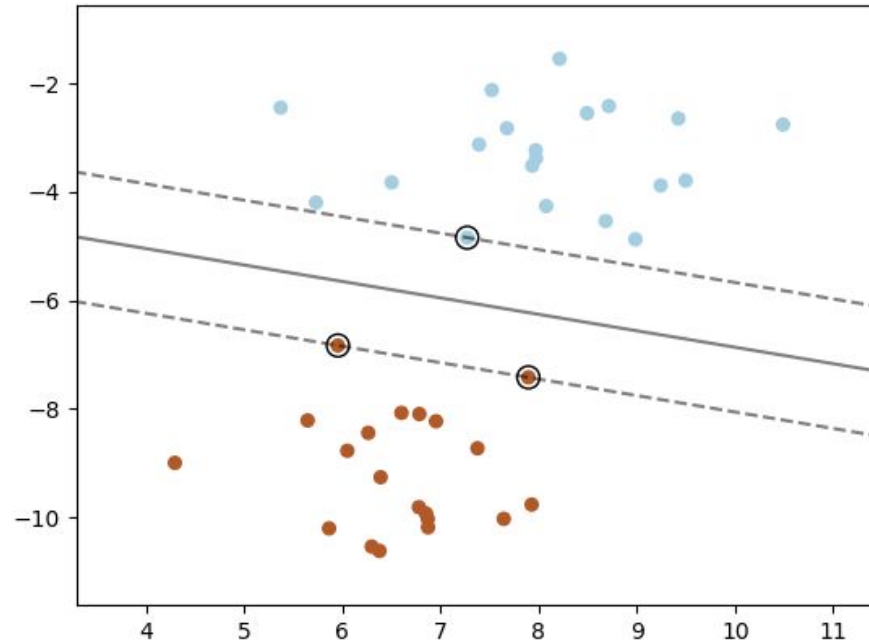
$$\hat{y} = \arg \max_y P(y) \prod_{i=1}^n P(x_i \mid y)$$

Floresta Aleatória (RF)



Fonte: Adaptada de IBM (2023).

Máquina de Vetor de Suporte (SVM)



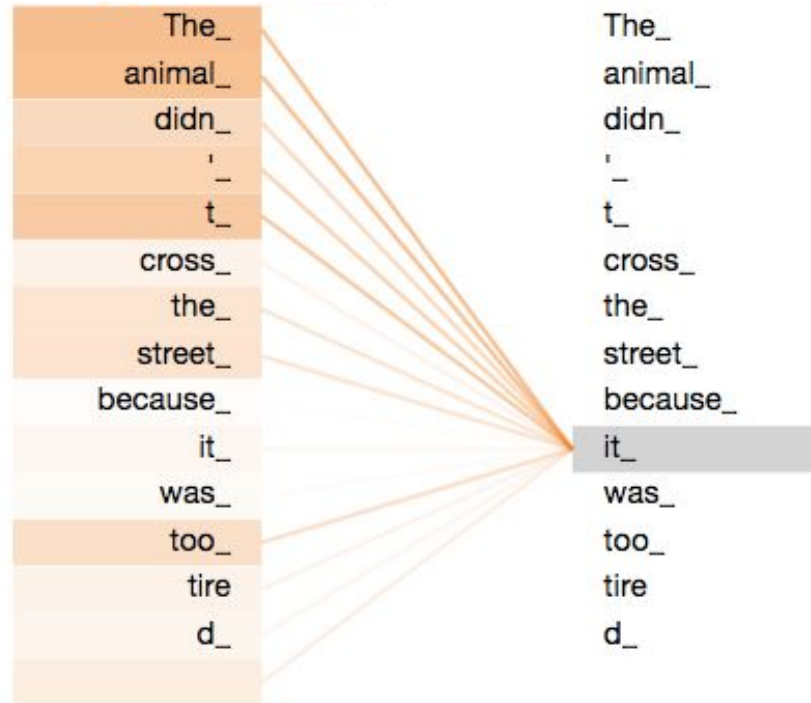
Fonte: Buitinck et al. (2013).

Mecanismo de Atenção

The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .
The FBI is chasing a criminal on the run .

Fonte: Cheng et al. (2016)

Transformers e Self-Attention



Fonte: Alammr (2020).

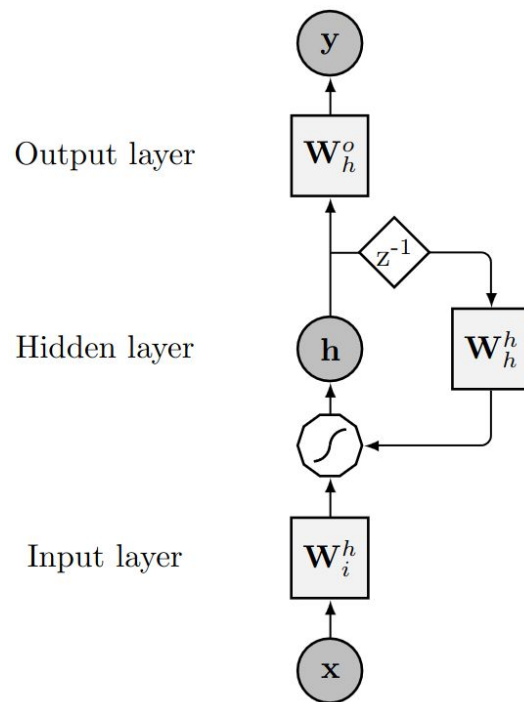
BERT

- *Bidirectional Encoder Representations from Transformers;*
- Pré-treinamento:
 - Modelo de Linguagem Mascarada (MLM);
 - *Next Sentence Prediction* (NSP);

BERTimbau

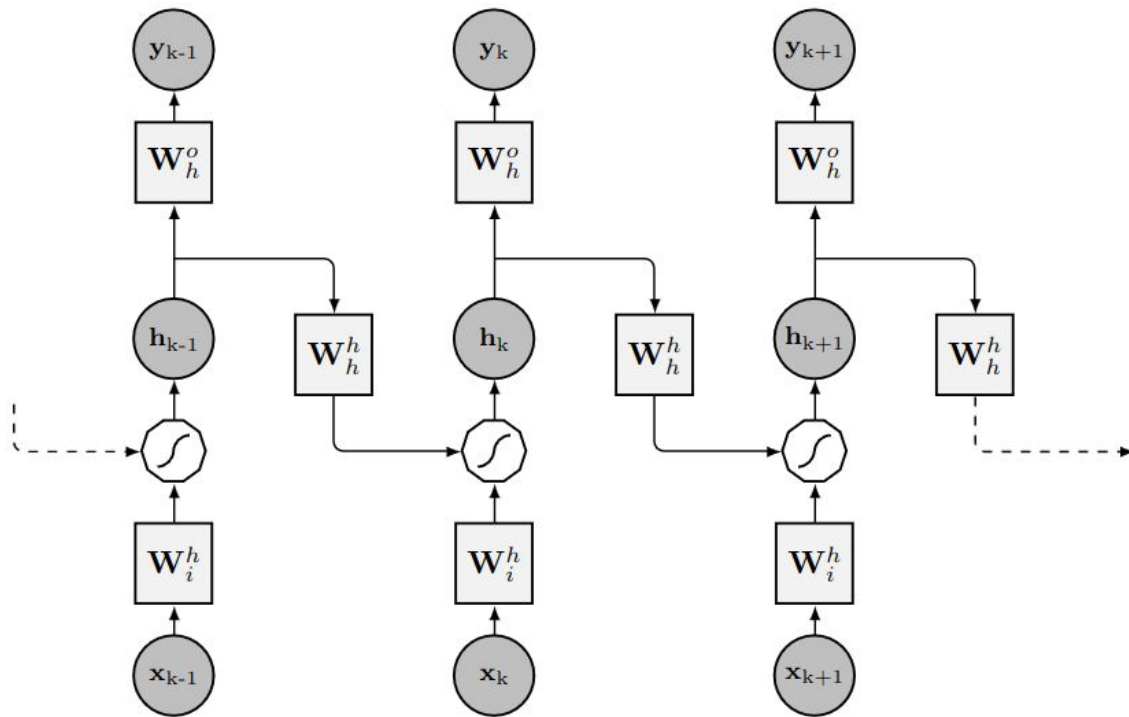
- Modelos monolíngues obtêm melhor desempenho que o BERT multilíngue, vantagens dos modelos monolíngues pré-treinados em dados de vários domínios;
- Treinado com o *Brazilian Portuguese Web as Corpus*, contendo 3.53 milhões de documentos.

RNN



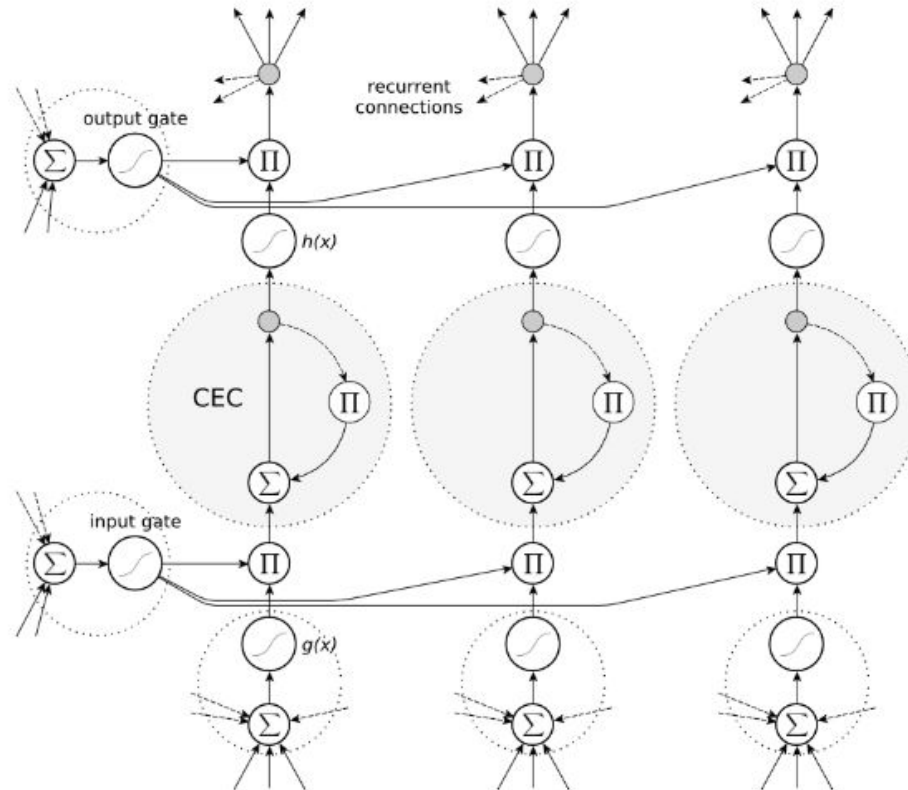
Fonte: Bianchi et al. (2017).

RNN



Fonte: Bianchi et al. (2017).

Long Short-Term Memory



Fonte: Staudemeyer e Morris (2019)

Metodologia e Resultados



Base de Dados

- Por Sadeghi, Bidgoly e Amirkhani (2022), com licença *Creative Commons Attribution 4.0*, permitindo o compartilhamento e a adaptação por atribuição;
- As amostras são preparadas em duas etapas:
 - Na primeira etapa, as amostras existentes do site <https://www.politifact.com/> foram agrupadas utilizando a API até 26 de abril de 2020;
 - No segundo passo, é efetuado um pré-processamento nos dados recolhidos para remover as partes do texto que representam o rótulo da veracidade das notícias.
- As amostras foram divididas em 15.212 para o treinamento, 1.058 para a validação e 1.054 para o teste. As classes destes dados são "*real*" e "*fake*".
- 80% para o treinamento, 10% para validação e 10% para teste.



Ron DeSantis in 2018 backed a bill drafting a path for Puerto Rico statehood, but did not author it

A TV ad in Iowa from a political action committee supporting former President Donald Trump tells voters that Florida Gov. Ron DeSantis has "sold out" conservatives.

"Liberals have a plan to make Puerto Rico a state, adding two Democrats to the Senate. And Ron DeSantis sided with the liberals' power play," [MAGA Inc.'s ad said](#). "DeSantis actually sponsored the bill to make Puerto Rico a state. With more power, liberals can pack the courts, pass more reckless spending, ban guns and give amnesty to illegal aliens."



A MAGA Inc. ad said DeSantis "actually sponsored the bill to make Puerto Rico a state."

DeSantis co-sponsored a 2018 bill that outlined steps that Congress could take to make Puerto Rico a state by 2021. He was not the bill's original sponsor or author — he was one of dozens of co-sponsors from both parties. The legislation outlined several steps a task force needed to take before statehood could occur.

DeSantis' current stance on this issue is unclear.

3. Label

We rate this claim Mostly True.

MAGA Inc

stated on October 28, 2023 in a TV ad in Iowa.

1. Statement

Gov. Ron DeSantis
"actually sponsored the
bill to make Puerto Rico a
state."



NATIONAL CONGRESS FLORIDA MAGA INC



Republican presidential candidate Ron DeSantis walks in the July 4th parade, July 4, 2023, in Merrimack, N.H. (AP)

Tradução com Googletrans

- Algumas vezes há inconsistências na tradução para o português brasileiro, sendo utilizada a grafia de outros países, como no uso da palavra “económica” em vez de “econômica” e “actual” em vez de “atual”;

Quantidade de Caracteres e Palavras

Rótulo	Análise	Média
Real	Caracteres nas premissas	3154,3997
	Palavras nas premissas	512,3448
	Caracteres nas hipóteses	125,8918
	Palavras nas hipóteses	20,8673
Falsa	Caracteres nas premissas	3063,3925
	Palavras nas premissas	497,2056
	Caracteres nas hipóteses	121,3452
	Palavras nas hipóteses	19,9495

Fonte: Elaborado pela autora.

Palavras

Word cloud com as 1.000 palavras mais utilizadas nos textos das declarações falsas



Fonte: Elaborada pela autora.

Word cloud com as 1.000 palavras mais utilizadas nos textos das declarações verdadeiras



Fonte: Elaborada pela autora.

Resultados

Resultados obtidos por Sadeghi, Bidgoly e Amirkhani (2022) sem o uso de premissas:

Modelo	GloVe	BERT
NB	0,6923	0,6691
RF	0,6520	0,6567
LR	0,6850	0,6949
SVM	0,7001	0,6766
BiLSTM	0,6170	0,7514

Fonte: Sadeghi, Bidgoly e Amirkhani (2022).

Resultados

Resultados obtidos por Sadeghi, Bidgoly e Amirkhani (2022) com o uso de premissas:

Modelo	GloVe	BERT
NB	0,6778	0,6623
RF	0,6666	0,7132
LR	0,7351	0,8007
SVM	0,7322	0,7607
BiLSTM	0,8512	0,8548

Fonte: Sadeghi, Bidgoly e Amirkhani (2022).

Resultados

Resultados obtidos sem o uso de premissas:

Modelo	Representação das Palavras	Macro-averaged F1 score
NB	BoW e TF-IDF	0,7162
RF	BoW e TF-IDF	0,6599
LR	BoW e TF-IDF	0,7125
SVM	BoW e TF-IDF	0,6885
BiLSTM	BERTimbau	0,7205

Fonte: Elaborado pela autora.

Resultados

Resultados obtidos com o uso de premissas:

Modelo	Representação das Palavras	Macro-averaged F1 score
NB	BoW e TF-IDF	0,8102
RF	BoW e TF-IDF	0,7732
LR	BoW e TF-IDF	0,8309
SVM	BoW e TF-IDF	0,8234
BiLSTM	BERTimbau	0,8199

Fonte: Elaborado pela autora.

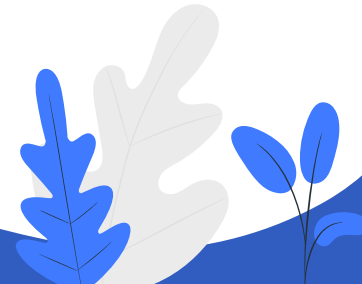


Conclusão

Conclusão


- Corroborou-se melhores resultados na língua portuguesa com o uso de premissas;
- Trabalhos futuros:
 - Em outras línguas;
 - Além disso, no caso do BERTimbau, uma possível forma de aumentar sua pontuação é contornar o problema do limite de 512 tokens, por exemplo, passando cada parágrafo para o seu decodificar e concatenando os embeddings para cada amostra.

Referências






Referências

- ALAMMAR, JAY. Illustrated: Self-Attention. **Medium**, 2020. Disponível em: <https://jalammar.github.io/illustrated-transformer/>. Acesso em: 14 jun. 2023;
 - BIANCHI, F. M.; MAIORINO, E.; KAMPFFMEYER, M.; RIZZI, A.; JENSSEN, R. Recurrent neural network architectures. In: . [S.l.: s.n.], 2017. p. 23–29. ISBN 978-3-319-70337-4.
 - BUITINCK, L.; LOUPPE, G.; BLONDEL, M.; PEDREGOSA, F.; MUELLER, A.; GRISEL, O.; NICULAE, V.; PRETTENHOFER, P.; GRAMFORT, A.; GROBLER, J.; LAYTON, R.; VANDERPLAS, J.; JOLY, A.; HOLT, B.; VAROQUAUX, G. API design for machine learning software: experiences from the scikit-learn project. In: ECML PKDD Workshop: Languages for Data Mining and Machine Learning. [S.l.: s.n.], 2013. p. 108–122.
 - CHENG, Jianpeng; DONG, Li; LAPATA, Mirella. Long Short-Term Memory-Networks for Machine Reading. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Anais eletrônicos [...] Association for Computational Linguistics: Austin, 2016. p. 551–561. Disponível em: <https://aclanthology.org/D16-1053/>. Acesso em: 18 jun. 2023;
 - HE, Pengcheng; LIU, Xiaodong; GAO, Jianfeng; CHEN, Weizhu. Microsoft DeBERTa surpasses human performance on the SuperGLUE benchmark. **Microsoft**, 2020. Disponível em: <https://www.microsoft.com/en-us/research/blog/microsoft-deberta-surpasses-human-performance-on-the-superglue-benchmark/>. Acesso em: 18 jun. 2023.
 - IBM. What is random forest? 2023. Disponível em: <https://www.ibm.com/topics/random-forest>. Acesso em: 21 de outubro de 2023.
- 



Referências

- SADEGHI, Fariba; BIDGOLY, Amir Jalaly; AMIRKHANI, Hossein. FNID: Fake News Inference Dataset. **IEEE**, 2022. Disponível em: <https://ieee-dataport.org/open-access/fnid-fake-news-inference-dataset>. Acesso em: 18 jun. 2023.
 - SHERMAN, Amy. Ron DeSantis in 2018 backed a bill drafting a path for Puerto Rico statehood, but did not author it. **PolitiFact**, 2023. Disponível em: <https://www.politifact.com/factchecks/2023/nov/01/maga-inc/ron-desantis-in-2018-backed-a-bill-drafting-a-path/>. Acesso em: 10 nov.. 2023.
 - STAUDEMEYER, R. C.; MORRIS, E. R. Understanding LSTM - a tutorial into long short-term memory recurrent neural networks. CoRR, abs/1909.09586, 2019.
- 
- 