

STALLA: Um Framework para Análise de Fontes Abertas durante a Pandemia do Covid-19

Alexandre C. Godinho¹, Cristiano M. Nunes², Rodolfo I. Meneguette³,
Vinicius P. Gonçalves¹, Marlos André M. S. de Oliveira⁴,
José C. M. Oliveira⁴, Geraldo P. Rocha Filho⁴

¹Universidade de Brasília – UnB – Brasília – DF – Brasil

²Exército Brasileiro – EB – Brasil

³Universidade de São Paulo – USP – São Carlos – SP – Brasil

⁴Universidade Estadual do Sudoeste da Bahia – UESB – Vitória da Conquista – BA – Brasil

alexandre.godinho@aluno.unb.br, nunes.cristiano@eb.mil.br

meneguette@icmc.usp.br, vpgvinicius@unb.br

{marlos, jcarlos, geraldo.rocha}@uesb.edu.br

Abstract. *The spread of social networks has resulted in an increase in the distribution of disinformation campaigns, which put national democratic stability at risk, becoming an unfavorable element for the intelligence knowledge production. In order to mitigate this bottleneck, the STALLA framework was proposed for the collection, treatment, automated labeling and analysis of information, providing greater efficiency in knowledge production. Thus, the study has as scope the Covid-19 pandemic, from data collected from short texts (tweets), in the Portuguese language, from the social network Twitter. Considering the related works, Recurrent Neural Networks (RNN) present themselves as the most suitable for textual analysis. Based on this premise, the performance of STALLA was analyzed by comparing the implementations of LSTM and BiLSTM networks, resulting in an accuracy of approximately 70%.*

Resumo. *A expansão das redes sociais resultou em um aumento na distribuição de campanhas de desinformação, que colocam em risco a estabilidade democrática nacional, tornando-se um elemento desfavorável para a produção do conhecimento de Inteligência. Com o objetivo de mitigar este óbice, foi proposto o framework STALLA para coleta, tratamento, rotulação automatizada e análise de informações, proporcionando maior eficiência na produção do conhecimento. Assim, o estudo tem por escopo a pandemia do Covid-19, a partir de dados coletados de textos curtos (tweets), no idioma português, da rede social Twitter. Considerando-se os trabalhos correlatos, as Redes Neurais Recorrentes (RNN) apresentam-se como as mais vocacionadas para análises textuais. A partir dessa premissa, o desempenho do STALLA foi analisado comparando-se as implementações das redes LSTM e BiLSTM, resultando em uma acurácia de aproximadamente 70%.*

1. Introdução

A rápida expansão das redes sociais [Oshikawa 2020], associada à popularidade dos smartphones [Silva 2020], permite que uma enorme quantidade de informação seja publicada diariamente. Nesse contexto, os usuários de redes sociais buscam informações nessas plataformas, deixando para um segundo plano a imprensa. No entanto, a validação de veracidade dessas postagens é algo extremamente árduo, tendo em vista que é intencionalmente publicada para enganar as pessoas [Lazer 2018].

Assim, altera-se a percepção da sociedade sobre diversas temáticas, o que contribui para a divulgação de *fake news* [Bechmann 2018] e campanhas de desinformação, provocando ameaças aos estados nacionais democráticos, bem como à segurança pública nesses países [Kang 2016]. A velocidade de propagação é o mais alarmante [Torabi Asr 2019], sendo inegável que as redes sociais e as aplicações de mensagens instantâneas intensifiquem a divulgação de conteúdo enganoso [Vosoughi 2018].

Diante desse cenário, surgiram esforços para mitigar o crescimento das campanhas de desinformação [Silva 2020]. No entanto, em regra, as agências de *fact-checking* utilizam rotulagem manual, realizada por analistas especializados, o que é um procedimento caro, demorado e pode ser contaminado pelo viés cognitivo do profissional que o realiza [Oshikawa 2020]. Dessa forma, a comunidade de pesquisa identificou essa lacuna e vem desenvolvendo estudos com o objetivo principal de reduzir tempo e esforço humano.

Nesse contexto, pesquisas acadêmicas vêm contribuindo para: (i) identificar a velocidade de propagação das *fakes news* e campanhas de desinformação; (ii) analisar o comportamento dos usuários de redes sociais; (iii) propor frameworks capazes de verificar de forma automática a veracidade e a relevância dos conteúdos publicados em linguagem escrita; e (iv) explorar as características linguísticas que possibilitem identificar conteúdos enganosos [Silva 2020, Weigang et al. 2022].

Outrossim, o desenvolvimento de soluções tecnológicas, baseadas em quantidades massivas de dados e Inteligência Artificial, enseja oportunidades para a atividade de Inteligência, como por exemplo: a automação da coleta de dados, a redução do tempo de processamento na análise de estruturas de dados complexas e o refinamento dos resultados para apresentar os principais pontos que conduzam a uma tomada de decisão efetiva por um gestor público ou o responsável [Alves 2018, Brasil 2017, Enamoto et al. 2021, Enamoto et al. 2022].

Diante do exposto, a pesquisa em questão, por intermédio da implantação de *weak supervision* e *Recurrent Neural Network with Bidirectional Long Short-Term Memory* (RNN-BiLSTM), possui os seguintes objetivos precípuos: identificar as informações que geram maior engajamento e possuem maior potencial para fomentar campanhas de desinformação, bem como mitigar a carência de *datasets* temáticos, através da supervisão fraca. Para isso foi proposto o *framework* STALLA para atenuar os efeitos da subjetividade e conferir maior assertividade para a produção de Inteligência.

O trabalho está organizado da seguinte forma. Na Seção 2, são apresentados trabalhos relacionados. Na Seção 3, são apresentados os detalhes metodológicos para a produção do conhecimento de inteligência. Na Seção 4, é apresentada a modelagem do STALLA, enquanto na Seção 5 é apresentado o cenário para validar o STALLA. Por fim, na Seção 6, são apresentadas as conclusões obtidas e sugestões para trabalhos futuros.

2. Trabalhos Relacionados

No âmbito das pesquisas mencionadas, a disponibilidade de conjuntos de dados rotulados pode ser considerado um dos principais óbices para a automatização da detecção de notícias falsas. Existem alguns conjuntos de dados a respeito de *fake news* e campanhas de desinformação, mas no idioma português, o trabalho ainda é incipiente, havendo a oportunidade de desenvolver um conjunto de dados para treinar algoritmos de machine learning.

A partir desse desafio, foi empregado no presente estudo o *framework Snorkel*, [Ratner 2017], que permite a construção de conjuntos de dados de treinamento de forma programática, mitigando os custos de empregar recursos especializados, como no caso de jornalistas atuando como *fact checkers*.

Diversas técnicas de aprendizado de máquina vêm sendo empregadas para a detecção de campanhas de desinformação. Os estudos para o processo de detecção de *fake news* são baseados em pesquisas teóricas [Duran 2010, Hauch 2015] e práticas [Appling 2015, Pérez-Rosas 2015, Rubin 2016], integrando uma grande variedade de disciplinas, como a linguística e a ciência da computação [Torabi Asr 2019].

Os sistemas de detecção automática de *fake news* são baseados em métodos de classificação de textos [Torabi Asr 2019] e apresentam vantagens sobre o processo realizado de forma manual, pois os julgadores humanos são propensos a preconceitos. Ademais, julgamentos que dependam de várias fontes de informação, como vídeos, áudios e outras mídias podem sobrecarregar o julgador e levar a atrasos e erros.

Embora os sistemas de detecção automática não possam ser empregados de forma completamente independente, eles podem auxiliar os especialistas humanos nos procedimentos de *fact-checking* [Torabi Asr 2019]. Nesse aspecto, a identificação de conteúdo relevante permite que a classificação de texto aponte para alegações exageradas, linguagem excessivamente emocional ou um estilo incomum nas principais fontes de notícias. Desta forma, é possível a detecção de *fake news* com base no estilo, no contexto e no conhecimento de determinado assunto.

A detecção automatizada de campanhas de desinformação pode ser conceituada como a tarefa de avaliar a veracidade das afirmações nas notícias [Oshikawa 2020], empregando diversas técnicas de Inteligência Artificial, dentre elas *Machine Learning*, *Data Mining* e *Natural Language Processing*.

Nesse escopo, destacam-se as tarefas de classificação de rótulos multiclasse (falso, verdadeiro, parcialmente falso, parcialmente verdadeiro) [Rashkin 2017] e regressão [Nakashole 2014]. Uma das condições para que os sistemas classificadores de notícias falsas alcancem bons desempenhos é que existam dados rotulados suficientes [Oshikawa 2020], sendo mais comuns os modelos de *Machine Learning* que empregam métodos supervisionados de aprendizagem.

3. Produção do Conhecimento de Inteligência

3.1. Visão Geral

As principais agências de inteligência dos Estados nacionais utilizam uma metodologia [US 2014] que relaciona dados, informação e produção do conhecimento (Figura 1). Essa

metodologia permite identificar as tarefas precípua para a produção do conhecimento de inteligência, contribuindo para a modelagem do framework proposto nesse trabalho.

A metodologia destaca a importância do emprego de grandes quantidades de dados para o processamento e a obtenção de informações significativas, visando a produção do conhecimento e, conseqüentemente, o assessoramento do poder decisório nacional. Nesse sentido, foram levantados os procedimentos técnicos que representam as etapas de coleta; processamento e exploração; e análise e produção do conhecimento, afim de replicar o conceito metodológico com os dados oriundos de fontes abertas.

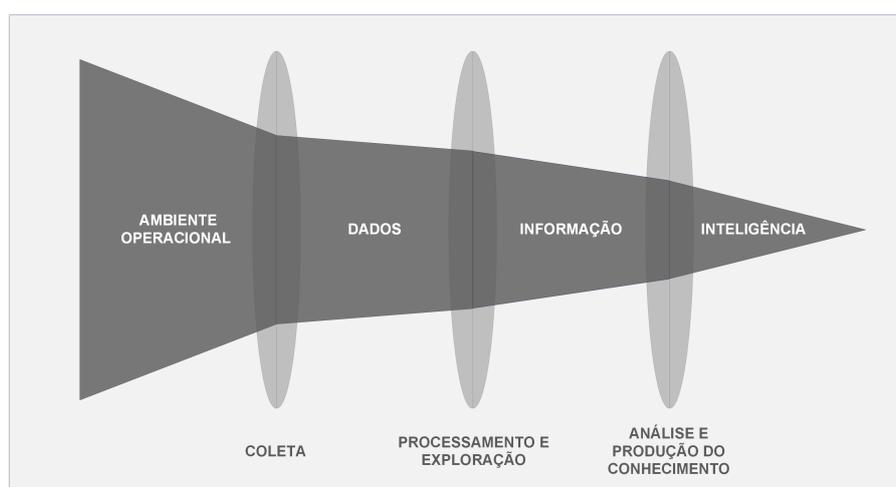


Figura 1. Relação entre dados, informação e produção do conhecimento

3.2. Coleta

Na fase de coleta, as agências de inteligência utilizam as redes sociais como fontes para a obtenção de um volume massivo de dados abertos, sendo necessário a avaliação acerca do grau de veracidade do conteúdo e confiabilidade da fonte. No entanto, devido à simplicidade e ao baixo custo de obtenção, torna-se congruente o emprego de técnicas para a produção do conhecimento a partir dessas fontes.

3.3. Processamento e Exploração

Durante a fase de processamento, realiza-se a carga dos dados brutos obtidos das fontes abertas em um banco de dados temporário para preparação, limpeza, tabulação e demais transformações necessárias. Nessa fase, também, são elaboradas as funções de rotulação, tendo por base o conhecimento de especialistas na temática a ser analisada. Com base nessas funções, desenvolve-se um *dataset* temático orientado ao estudo de caso proposto.

A exploração dos dados é uma tarefa acessória, cujo principal objetivo é enriquecer o quadro de referência do analista de inteligência para conceber as funções de rotulação, o mais assertivamente possível. Para isso, emprega-se técnicas de *analytics* com o propósito de visualizar os dados e as suas métricas, como também deduzir os correlacionamentos entre as entidades, resultando em diagramas, gráficos e nuvens de palavras.

3.4. Análise e Produção do Conhecimento

É sabido que o trabalho de produção do conhecimento de inteligência é muito dispendioso, uma vez que demanda um estudo analítico de grandes quantidades de informação, fenômeno conhecido como “*information overload*”, além do quadro de referência do analista de inteligência, para se obter o conhecimento de interesse [Alves 2018].

No intuito de mitigar esse óbice, buscou-se automatizar o processo de categorização de uma informação quanto a sua relevância. Assim, por intermédio de técnicas de aprendizagem de máquina, aplica-se o quadro de referência do analista, implementado a partir das funções de rotulação, para identificar dentre o excesso de informações, aquelas que possuem maior relevância para a produção do conhecimento.

4. Framework STALLA

O *framework* proposto no presente trabalho remete ao acrônimo do idioma inglês, *scraping, transforming, auto-labelling, learning and analysis* (Figura 2). O *framework* STALLA contempla em sua nomenclatura as etapas de procedimentos técnicos necessários para a obtenção desde a coleta dos dados até análise e produção do conhecimento, alcançando a extremidade do “funil” da inteligência.

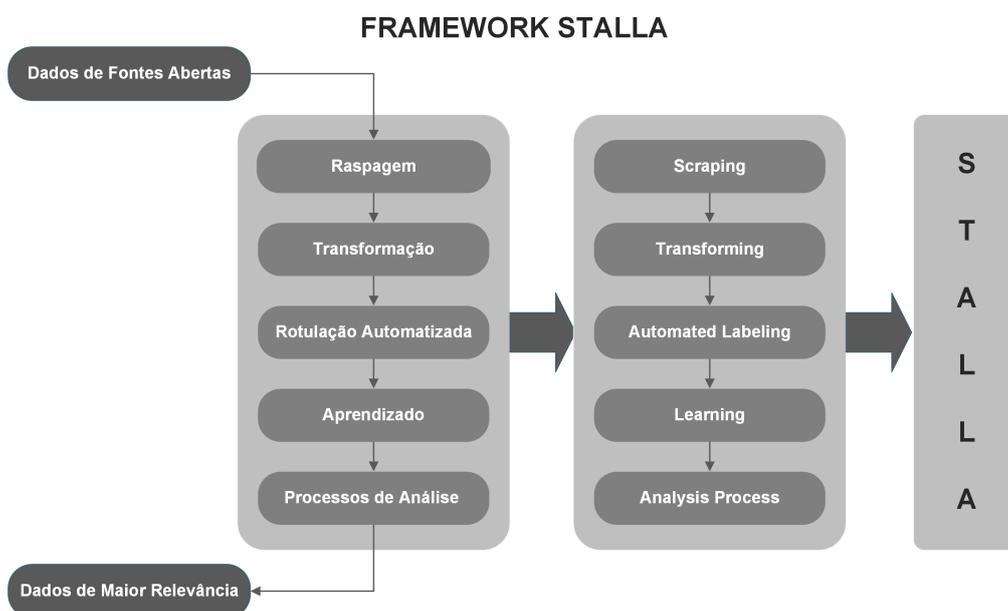


Figura 2. Framework STALLA

4.1. Raspagem de Dados

A raspagem de dados (ou *scraping* no idioma inglês) consiste na obtenção de informações renderizadas na página *html* quando acessada por um *browser*. A depender das tecnologias envolvidas, a complexidade da raspagem pode ser maior, pois alguns sites implementam certos obstáculos para execução do *scraping*. Em contrapartida, existem sites provedores de conteúdo que disponibilizam APIs para facilitar o consumo dos seus dados.

Como consolidação do trabalho experimental podemos identificar como o Framework STALLA é desenvolvido através de sua visão geral aplicada neste caso de estudo da pandemia do Covid-19, Figura 4.

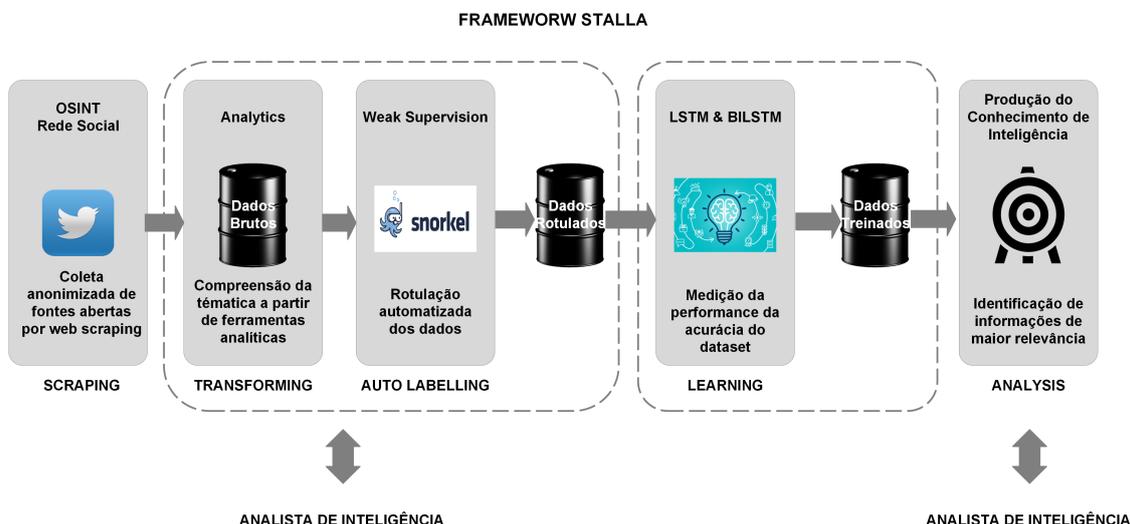


Figura 4. Visão geral das funções do FRAMEWORK STALLA

5. Experimentos

5.1. Configuração dos Experimentos

Os dados utilizados nessa pesquisa são provenientes de uma coleta na rede social Twitter, em virtude desse *microblog* possuir um número abundante de fontes de informação sobre a temática da Covid-19. O procedimento de coleta consiste no uso de web crawlers, instâncias de máquinas virtuais em nuvem, que executam um código para realização de *web scraping*.

A técnica de *web scraping* foi executada através de scripts desenvolvidos na linguagem de programação *Python*, que utilizam funções das bibliotecas públicas *Snsrape* e *Requests* [Snsrape 2023]. Na Tabela 1 encontram-se as características da referida coleta.

Tabela 1. Qualificação do processo de obtenção de dados

Fonte de OSINT	Rede Social Twitter
Idioma	Português
Quadro de Referência	Tweets contendo a palavra “covid” (incluindo variações de maiúsculas e minúsculas)
Período de coleta	Maior de 2021 a Maior de 2022
Amostragem	Aproximadamente 125k <i>tweets</i> /mês
Qtd de web crawlers	02 com 25 <i>threads</i> de execução cada
Tempo de execução	120 min

Campos coletados	nome de usuário, quantidade de seguidores, localização, quantidade de curtidas, quantidade de retweets, quantidade de réplicas, data da publicação, dispositivo de origem e corpo textual
Qtd de total tweets coletados	1.525.775
MD5 Hash	c6b5ebc0cb30974fe7eeb8332adee612
Disponível em	https://shorturl.at/dGKR8

Os dados coletados foram carregados em um *dataframe Pandas*, sendo cada coluna preenchida por um campo de interesse, a saber: nome de usuário, quantidade de seguidores, local de origem da postagem (Figura 5), quantidade de *likes* (Tabela 2), quantidade de réplicas (Tabela 3), quantidade de *retweets* (Tabela 4), data da postagem, dispositivo de origem da postagem e corpo textual do tweet.

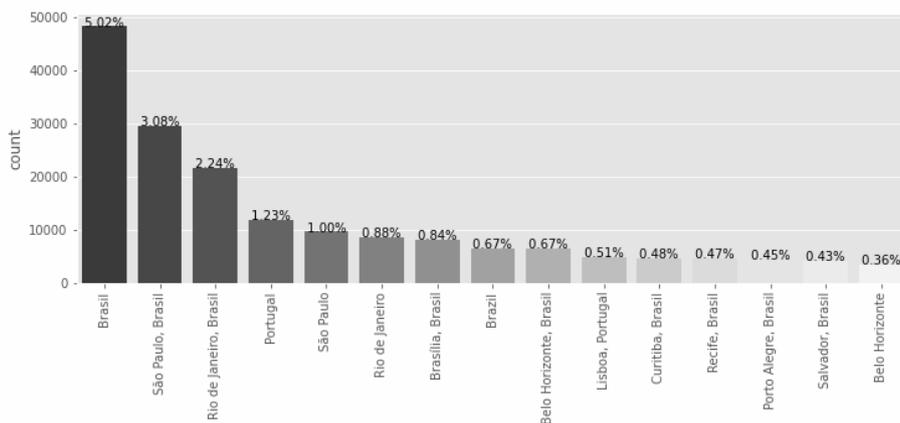


Figura 5. Local de origem dos tweets

Ainda sobre os dados referentes à pandemia da Covid-19, obtidos da plataforma Twitter, destaca-se que na maioria das vezes essas informações não estão disponíveis no tweet, uma vez que isso depende das políticas de privacidade selecionadas pelo usuário.

Outro ponto a ressaltar, é a politização da temática relacionada à pandemia do Covid-19, o que pode ser observado pelo emprego de bibliotecas *Python WordCloud* e *Matplotlib* que resultaram em visualizações que contribuem para ampliar o quadro de referência do analista de inteligência, favorecendo a elaboração das funções de rotulação.

Tabela 2. Top 3 de tweets com o maior número de likes

Usuário	Nr Likes	Corpo textual do tweet
Phenrique1201	177025	“enfim zeramos o cti covid”
aliceeleroy	170440	“vim fazer teste de covid e a moca ”com qual cor vc se identifica?”eu: verde caras, as opcoes eram branca/parda/negra”

Tabela 2. Top 3 de tweets com o maior número de likes

Usuário	Nr Likes	Corpo textual do tweet
predonisona	152959	“e agora sera que esses sintomas sao de rinite sinusite h1n1 h3n2 covid h2o ou hb20 hbo max”

Tabela 3. Top 3 de tweets com o maior número de réplicas

Usuário	Nr Réplicas	Corpo textual do tweet
pdrbnt	5391	“vc conseguiu tomar a vacina antes de contrair covid?”
gen_helena	4224	“impressionante a parcialidade da cnn ao comentar a cpi da covid. algo constrangedor. jornalista mona lisa, dany lima e alguns colegas deviam se unir aos seis senadores da oposicao para se postarem contra o governo. garanto que a audiencia ja vem punindo essa lamentavel postura.”
RomeuZema	3571	“entre os internados com covid em minas hoje, os que nao se vacinaram sao proporcionalmente tres vezes mais do que aqueles que tomaram todas as doses. o problema nao esta na vacina e sim na desinformacao. vacinas salvam vidas! vacine.”

Tabela 4. Top 3 de tweets com o maior número de retweets

Usuário	Nr Retweets	Corpo textual do tweet
predonisona	36203	“e agora sera que esses sintomas sao de rinite sinusite h1n1 h3n2 covid h2o ou hb20 hbo max”
KriskaCarvalho	11862	“ele imitou uma pessoa morrendo sem ar com covid, ele riu, ele debochou de alguem sem ar num leito... morreram quase 700 mil ou mais... ele riu, nunca esqueco!”
vozdareisist	11714	“se fizermos um minuto de silencio para cada pessoa morta por covid-19, ficaríamos calados por oito anos. imagem da manifestacao no rio de janeiro.”

Para o desenvolvimento do *dataset* sobre os tweets de maior valor para o analista foi utilizado o *Snorkel Framework*, biblioteca *Python* que permite codificar funções de rotulação para automatizar este processo, no quadro a seguir é apresentado um exemplo de algumas funções de rotulação utilizadas.

```
from snorkel.labeling import labeling_function

@labeling_function()
def lf_14(x):
```

```

    return HIGHVALUE if x.Followers > 100 else ABSTAIN

@labeling_function()
def lf_19(x):
    return HIGHVALUE if (x.No_Likes > 5) and (x.No_RT > 15) else ABSTAIN

@labeling_function()
def lf_24(x):
    return LOWVALUE if re.search(r"Android|iPhone", x.Source_device, flags=
                                re.I) and (x.Followers < 20) else
                                ABSTAIN

```

O modelo de aprendizado de máquina para rotulação adotado foi o *Majority Vote*, que é encarregado de reunir os resultados das funções de categorização em um único rótulo. Para isso, foram definidos os seguintes parâmetros das funções de rotulação: a quantidade de seguidores dos autores dos tweets, a quantidade de curtidas dos tweets, a quantidade de retweets de cada tweet, a quantidade de réplicas de cada tweet, o dispositivo ou plataforma de publicação, o local de publicação, os pronomes do corpo textual e o quadro de referência do analista.

Para avaliar o modelo proposto foram adotadas as arquiteturas *Long Short-Term Memory* (LSTM) e *Bidirectional Long Short-Term Memory* (Bi-LSTM), comparando-se o resultado proveniente dessas duas implementações. Para o sequenciamento de dados e a divisão do processamento, os dados foram convertidos em vetores tokenizados, com valores de rótulos representativos. A estrutura da divisão do *shape* de treinamento com a biblioteca *scikit-learn* foi a seguinte:

$$X_{train} : 360018; X_{test} : 120006; Y_{train} : 360018; eY_{test} : 120006$$

Como preparação dos dados, o corpo de texto dos tweets foram tratados com a biblioteca de processamento de linguagem natural *Gensim* que emprega modelos acadêmicos para realizar a tokenização das palavras. Na sequência, foram criados arrays através da biblioteca *Numpy* para configurar os dados no formato adequado para emprego da rede neural.

Com o *dataset* e os corpos dos tweets preparados em *arrays*, foi empregada a biblioteca *Keras*, que é a implementação de uma interface para customização e desenvolvimento de redes neurais artificiais em *Python*. As características definidas para rede por parâmetros e por configurações de hardware estão descritas na Tabela 5.

Tabela 5. Parâmetros de treinamento

Variáveis do treinamento	Descrição
Plataforma	Jupyter Notebook/ Python 3.9
GPU	NVIDIA GeForce GTX 1060
Optimizer	RMS prop
Loss	Categorical cross-entropy
Épocas	10
Ativação	Softmax

Nos experimentos executados foram configurados os seguintes parâmetros: o

”*Categorical Cross-entropia*”, como função de perda devido, a unificação dos rótulos verdadeiros em uma única unidade; o popular algoritmo de otimização adaptativo *RMS-prop*; a função de ativação *Softmax*, para normalização da saída final da rede e 10 (dez) épocas de treinamento.

5.2. Resultados e Discussão

A partir das configurações definidas, foram executados os treinamentos levando-se em conta os dois tipos de redes neurais, a LSTM e a Bi-LSTM. A diferença das implementações consiste na característica que na rede bidirecional o fluxo ocorre em ambas direções sendo capaz de utilizar informações de ambos os lados. Na Figura 6 pode-se notar uma pequena vantagem no desempenho da rede Bi-LSTM.

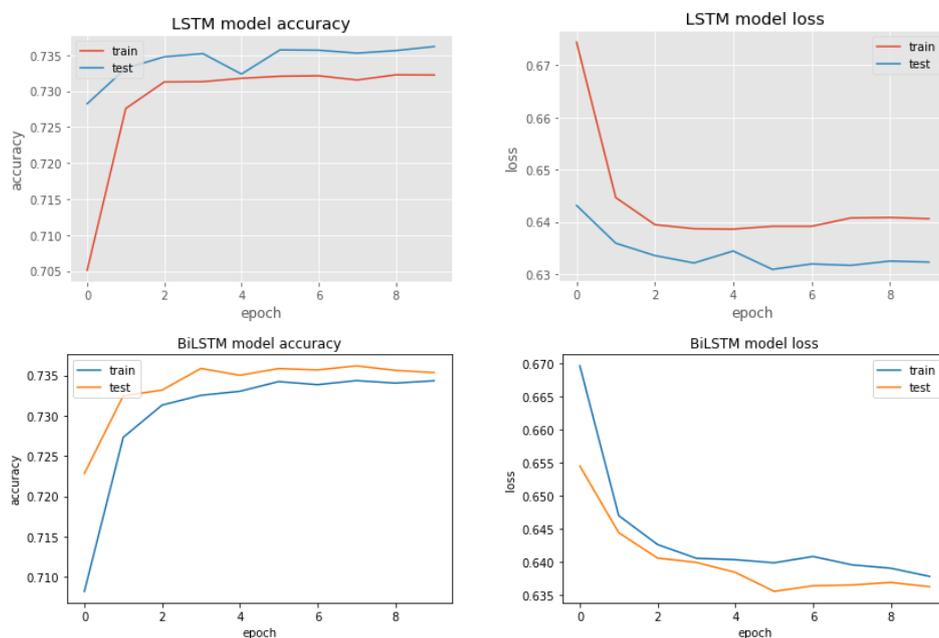


Figura 6. Resultado comparativo LSTM X Bi-LSTM

Os experimentos indicam ainda que com o aumento da quantidade de dados para treinamento, há uma evolução da acurácia do modelo Bi-LSTM, conforme pode ser constatado na Figura 7.

Diante do exposto, verifica-se que quantidade acima de 1 milhão de tweets como amostra e tendo por fundamento os critérios tomados para a rotulação do *dataset* de treinamento, foi possível prever o potencial de um tweet recém postado, enquadrar-se a padrões já identificados e, portanto, ser estatisticamente de maior relevância.

Para isso, foi adotado o *F1-score*, que é uma medida de performance da acurácia de um modelo sobre o *dataset*. Cabe destacar que quanto mais próximo de 1 melhor é o modelo.

$$Precisão = \frac{VerdadeirosPositivos}{VerdadeirosPositivos + FalsosPositivos}$$

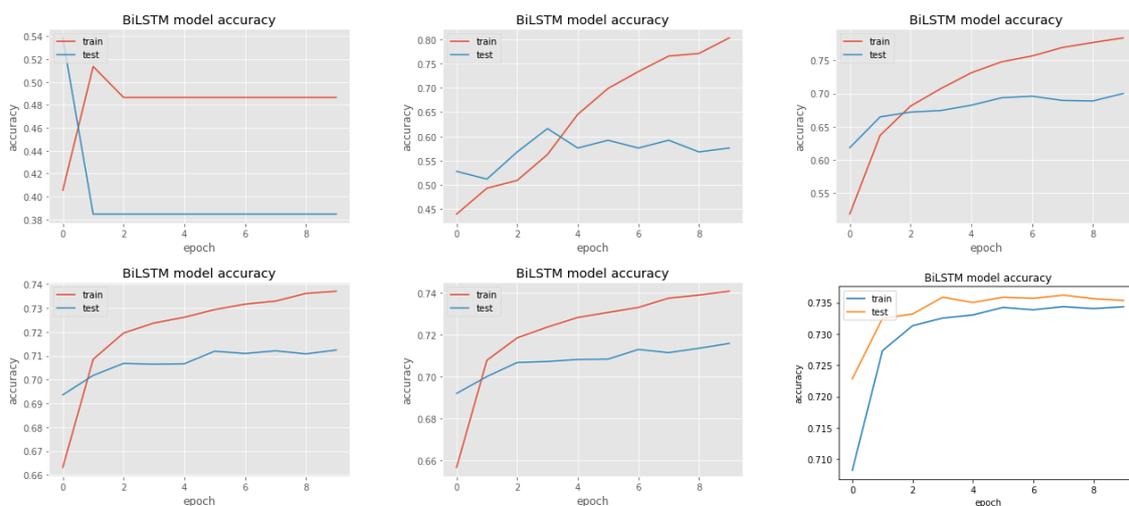


Figura 7. Evolução da acurácia frente à quantidade de dados

$$Recall = \frac{VerdadeirosPositivos}{VerdadeirosPositivos + FalsosNegativos}$$

$$F1 = \frac{2 * Precisão * Recall}{Precisão + Recall}$$

Ressalta-se que há contextos em que valores de *F1-score* podem ter interpretações muito específicas. Um exemplo clássico é o modelo de identificação de fraudes bancárias, no qual o valor de 0.95, pode ser um valor extremamente baixo, em função da finalidade a que se propõe, impactando em prejuízos financeiros inadmissíveis. No entanto, nesse estudo de caso, a avaliação da qualidade e da relevância das informações, é uma tarefa extremamente desafiadora, devido à enxurrada de dados produzidos nas fontes abertas, sem o compromisso de prover conhecimento verossímil. Nesta conjuntura, um resultado de *F1-score* superior a ½ representa um valor positivo na análise de conteúdo Tabela 6.

Tabela 6. Métricas de desempenho do Framework STALLA

-*-	Precision	Recall	F1-Score
Accuracy			0.74
Macro Avg	0.49	0.52	0.50
Weighted Avg	0.71	0.74	0.72

6. Conclusão

A identificação de uma informação que possui potencial para mobilizar usuários da plataforma, aproxima-se da finalidade precípua da Atividade de Inteligência, que é obter o conhecimento em momento oportuno para subsidiar a tomada de decisão pelo gestor responsável. Dessa maneira, a pesquisa resultou no Framework STALLA que permite alcançar uma maior produtividade na análise de informações de fontes abertas por parte do analista de inteligência.

Conclui-se que o modelo de predição treinado para detectar informações de maior relevância, no contexto da pandemia do Covid-19, atingiu aproximadamente 70% de acurácia, expondo a efetividade do framework STALLA na identificação de informações, potencialmente, mais relevantes, mesmo não se tratando de informações necessariamente verdadeiras. A acurácia de 70% é exclusiva para o caso de estudo que se busca aplicar uma interpretação de valor automatizada com viés de interesse do analista para dados desconhecidos. Em outras temáticas e dependendo da qualidade das funções de classificação é esperado que o valor da acurácia varie. A temática da Covid é uma temática complexa com grandes divergências, o valor de 70% seria considerado baixo em outros contextos.

Por fim, como trabalhos futuros, planeja-se desenvolver um mecanismo para realização de interpretação semântica com mais precisão, bem como testar o framework nos casos da Guerra da Ucrânia e das eleições presidenciais de 2022, realizadas no Brasil.

Agradecimentos: Geraldo P. Rocha Filho agradece à FAPESP (processo 2021/06210-3) e ao CNPq (304264/2022-9, PQ-2) por apoiarem seus projetos de pesquisas.

Referências

- Alves, P. M. M. R. (2018). O impacto do big data na atividade de inteligência. In *Revista Brasileira de Inteligência*. Brasília, DF, n. 13, p. 01-20, dez, 2018.
- Appling, Darren Scott; Briscoe, E. J. H. C. J. (2015). Discriminative models for predicting deception strategies. In *Proceedings of the 24th International Conference on World Wide Web*, pages 947–952.
- Bechmann, Anja; Nielbo, K. L. (2018). Are we exposed to the same “news” in the news feed? an empirical analysis of filter bubbles as information similarity for danish facebook users. In *Digital journalism*, v. 6, n. 8, pages 990–1002.
- Brasil (2017). Decreto nº 8.793, de 29 de junho de 2016. aprova a política nacional de inteligência. In *Diário Oficial da República Federativa do Brasil, Brasília, DF, Edição 241, 18 dez. 2017. Seção 1, p. 36-39.*
- Duran, N. D. e. a. (2010). The linguistic correlates of conversational deception: Comparing natural language processing technologies. In *Applied Psycholinguistics*, v. 31, n. 3, p. 439-462, 2010.
- Enamoto, L., Santos, A. R., Maia, R., Weigang, L., and Filho, G. P. R. (2022). Multi-label legal text classification with bilstm and attention. *International Journal of Computer Applications in Technology*, 68(4):369–378.
- Enamoto, L., Weigang, L., and Filho, G. P. R. (2021). Generic framework for multilingual short text categorization using convolutional neural network. *Multimedia Tools and Applications*, 80:13475–13490.
- Hauch, V. e. a. (2015). Are computers effective lie detectors? a meta-analysis of linguistic cues to deception. In *Personality and social psychology Review*, v. 19, n. 4, p. 307-342, 2015.
- Kang, Cecilia; Goldman, A. (2016). In washington pizzeria attack, fake news brought real guns. In *New York Times*, v. 5, 2016.

- Lazer, D. M. e. a. (2018). The science of fake news. In *Science*, v. 359, n. 6380, p. 1094-1096, 2018.
- Nakashole, Ndapandula; Mitchell, T. (2014). Language-aware truth assessment of fact candidates. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2014. p. 1009-1019.
- Oshikawa, Ray; Qian, J. W. W. Y. (2020). A survey on natural language processing for fake news detection. In *Conference on Language Resources and Evaluation (LREC 2020)*, pages 6086–6093.
- Pérez-Rosas, Verónica; Mihalcea, R. (2015). Experiments in open domain deception detection. In *Proceedings of the 2015 conference on empirical methods in natural language processing*. 2015. p. 1120-1125.
- Rashkin, H. e. a. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 conference on empirical methods in natural language processing*. 2017. p. 2931-2937.
- Ratner, A. J. e. a. (2017). Snorkel: Fast training set generation for information extraction. In *Proceedings of the 2017 ACM international conference on management of data*. 2017. p. 1683-1686.
- Rubin, V. L. e. a. (2016). Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection*. 2016. p. 7-17.
- Silva, R. M. e. a. (2020). Towards automatically filtering fake news in portuguese. In *Expert Systems with Applications*, v. 146, p. 113199, 2020.
- Snsrape (2023). Snsrape: A social networking service scraper in python. GitHub.
- Torabi Asr, Fatemeh; Taboada, M. (2019). Big data and quality data for fake news and misinformation detection. In *Big Data Society*, v. 6, n. 1, p. 2053951719843310, 2019.
- US (2014). Joint intelligence: Joint publication 2-0. Createspace Independent Pub.
- Vosoughi, Soroush; Roy, D. A. S. (2018). The spread of true and false news online. In *Science*, v. 359, n. 6380, p. 1146-1151, 2018.
- Weigang, L., Enamoto, L. M., Li, D. L., and Rocha Filho, G. P. (2022). New directions for artificial intelligence: human, machine, biological, and quantum intelligence. *Frontiers of Information Technology & Electronic Engineering*, 23(6):984–990.