

## **Protocolo da Revisão Sistemática da Literatura**

O objetivo desta revisão é realizar um levantamento da literatura relevante acerca do tema de detecção de fake news em mídias sociais. Este trabalho possui como meta investigar os modelos e métodos utilizados atualmente na detecção de fake news em mídias sociais, verificando se é uma notícia real ou não.

Para que o objetivo desta revisão seja cumprido, foi elaborado um protocolo de busca, para que seja levantado o máximo possível da literatura sobre o assunto. A seguir será descrito como o protocolo de busca foi elaborado.

### **PERGUNTA DE PESQUISA**

- **Pergunta:** Quais modelos são utilizadas para detectar se uma notícia é real ou não?
- **População:** Pesquisas na área de detecção falsas notícias (fake news) publicadas em bases de dados eletrônicas.
- **Intervenção:** Análise das pesquisas em andamento ou concluídas na área de falsas notícias (fake news).
- **Resultados:** Comparativo das pesquisas publicadas na área de detecção de falsas notícias.
- **Contexto:** Artigos publicados nas bases de dados eletrônicas: Science Direct e ACM.

### **BASES DE DADOS E STRINGS DE BUSCA**

A seguir serão descritas as strings de busca ou descrição da busca (quando as fontes de pesquisa não permitirem string de busca) que foram utilizadas em todas as fontes de pesquisa que foram consultadas nesta revisão sistemática. Cada fonte de pesquisa possui um mecanismo de busca próprio, por isso, as strings tiveram que ser adaptadas a cada fonte de pesquisa.

**Science Direct** - <https://www.sciencedirect.com/>

Find articles with these terms: fake news detection

Title, abstract or author-specified keywords: fake news

Year(s): 2008-2019

Article types: Review articles, research articles, data articles

**ACM** - <https://dl.acm.org/dl.cfm>

Select item from: The ACM Guide to Computing Literature

```
"query": { acmdlTitle:(+fake +news +detection) }
```

```
"filter": {"publicationYear":{ "gte":2008, "lte":2019 }},
```

```
{owners.owner=GUIDE}
```

## **CRITÉRIOS E PROCEDIMENTOS DE SELEÇÃO DOS TRABALHOS**

Após a construção e uso das strings de busca em cada fonte de pesquisa, devido a relevância do tema, centenas de trabalhos foram encontrados. Por este motivo, os trabalhos foram filtrados a partir dos seguintes critérios:

- Exclusão dos trabalhos repetidos, ou seja, que foram encontrados em mais de uma fonte de pesquisa;
- Pela análise do título do trabalho;
- Pela análise do resumo e da introdução do trabalho.

Os trabalhos foram avaliados e selecionados, principalmente em relação à qualidade e aderência ao tema da dissertação. Para que isso fosse possível, os trabalhos foram selecionados com base em quatro critérios de inclusão e dois critérios de exclusão. Os critérios de inclusão e exclusão são descritos a seguir.

### **Critérios de inclusão**

Os critérios de inclusão utilizados foram:

1. Trabalhos que apresentem um modelo de detecção para determinar se é uma notícia verdadeira ou não;
2. Trabalhos publicados entre 01/01/2008 até 07/11/2019;
3. Trabalhos que descrevam como a técnica foi utilizada para fazer a detecção;
4. Trabalhos publicados em língua inglesa ou portuguesa;

### **Critérios de Exclusão**

Os critérios de exclusão utilizados foram:

1. Trabalhos que não tenham foco na detecção de fake news. Por exemplo, trabalhos que foquem na previsão de fake news;
2. Trabalhos curtos (Resumos expandidos);

## **EXECUÇÃO DO PROTOCOLO DE REVISÃO SISTEMÁTICA DA LITERATURA**

As strings de busca foram realizadas em diversos sites, no entanto nem todas os sites apresentaram os resultados esperados. Inicialmente a busca foi realizada utilizando as palavras ‘data analysis’, entretanto os resultados estavam muito específicos na área de análise de dados.

Alguns testes com as strings de busca também foram realizados utilizando a palavra ‘prediction’, entretanto predição não é a alternativa correta quando se trata de verificação se uma notícia é verdadeira ou não, com isso as strings de busca finais foram realizadas utilizando as palavras ‘fake news’ e ‘detection’, pois o objetivo é detectar fake news e não prever.

Depois dos testes iniciais e a realização dos refinamentos nas strings de busca, foi iniciada a coleta e análise dos trabalhos. Os resultados iniciais das strings de busca podem ser visualizados abaixo:

Fonte de Busca	Total
Science Direct	28
ACM	32

**Tabela 1. Número de artigos encontrados**

A próxima etapa, seguindo os critérios e procedimentos de seleção dos trabalhos, foi a retirada dos trabalhos repetidos, entretanto como foram utilizados somente duas bases de dados, não houve nenhum trabalho repetido, como pode ser visualizado na tabela abaixo.

Fonte de Busca	Total
Science Direct	28
ACM	32

**Tabela 2. Número de artigos encontrados retirando repetidos**

A etapa seguinte consistiu da leitura do título de todos os trabalhos selecionados, verificando se estes se encaixavam nos critérios de inclusão ou em algum critério de exclusão. Os trabalhos ficaram distribuídos da seguinte forma:

Fonte de Busca	Total
Science Direct	10
ACM	21

**Tabela 3. Número de artigos a partir da análise dos títulos**

A etapa final envolveu a leitura dos resumos dos trabalhos restantes. Os trabalhos ficaram distribuídos da seguinte forma:

Fonte de Busca	Total
Science Direct	5
ACM	5

**Tabela 4. Número de artigos a partir da análise dos resumos**

Os 10 trabalhos restantes foram lidos, analisados e dispostos em um quadro no qual podem ser comparados seguindo os critérios descritos a seguir:

- Utilizou IA
- Qual técnica foi utilizada
- Variáveis de Entrada
- Tipo de Dado Classificado

- Utiliza Dados Reais
- Precisão de classificação

Os autores de cada artigo escolhido podem ser visualizados na tabela abaixo.

ID	Autor / Ano
1	Aldwairi, Monther; Alwahedi, Ali / 2018
2	Ko, Hoon; Hong, Jong Youl; Kim, Sangheon; Mesicek, Libor; Na, In Seop / 2019
3	Ozbay, Feyza Altunbey; Alatas, Bilal / 2019
4	Zhang, Chaowei; Gupta, Ashish; Kauten, Christian; Deokar, Amit V; Qin, Xiao / 2019
5	Jang, S Mo; Geng, Tieming; Li, Jo-Yun Queenie; Xia, Ruofan; Huang, Chin-Tser; Kim, Hwalbin; Tang, Jijun / 2019
6	Vedova, Marco Della; Tacchini, Eugenio; Moret, Stefano; Ballarin, Gabriele; DiPierro, Massimo; de Alfaro, Luca / 2018
7	Tschiatschek, Sebastia; Singla, Adis; Gomez Rodriguez, Manuel; Merchant, Arpit; Krause, Andreas / 2018
8	Gereme, Fantahun Bogale; Zhu, William / 2019
9	Qian, Feng; Gong, Chengyue; Sharma, Karishma; Liu, Yan / 2018
10	Khattar, Dhruv; Goud, Jaipal Singh; Gupta, Manish; Varma, Vasudeva / 2019

**Tabela 5. Lista de autores e ano de publicação**

Abaixo é apresentado uma tabela com a comparação entre cada um dos artigos selecionados.

ID	Usou IA?	Técnica Utilizada	Variáveis de Entrada	Tipo de Dado Classificado	Utiliza Dados Reais?	Precisão Classificação
1	Sim	Aprendizado de máquina com WEKA	Título começa com número?, título contém '?' e/ou '!', usuários deixaram o site depois de visitar?, todas as palavras são maiúsculas no título?, conteúdo não tem palavras do título?, título contém palavras-chave?	Páginas WEB, como: Facebook e Reddit.	Sim	99,40%
2	Não	Sistemas cognitivos baseado em backtracking	Utilização de palavras, principalmente palavras-chaves	Publicações da WEB	Sim	85,00%
3	Sim	23 algoritmos de IA, como: MLP, SGD, KLP	Confiança dos usuários, emoção social, pré - processamento dos dados definido filtrando os termos ou caracteres redundantes, como números, palavras de parada	Textos de mídias sociais	Sim	Variado de acordo com a rede
4	Não	Mecanismo de classificação baseado em tópicos, agrupando de notícias por tópicos	Medida de credibilidade para avaliar a autenticidade de uma notícia	Notícias	Sim	92.49%
5	Não	Árvore de evolução	Publicações do twitter, análise de notícias falsas populares	Falsos tweets relacionados as eleições presidenciais	Sim	
6	Sim	Aprendizado de máquina	Corpo das publicações, características dos usuários	Publicações do facebook e páginas	Sim	81,70%
7	Sim	Inferência bayesiana	Comportamento de sinalização dos verdadeiros usuários e usuários falsos	Notícias	Sim	
8	Sim	Deep learning	Características linguísticas e características dos usuários	Posts de mídias sociais	Sim	99,00%
9	Sim	Rede Neural Convolucional de Dois Níveis	Texto do artigo, artigos anteriores do usuário	Artigos	Sim	
10	Sim	Rede ponta a ponta, MVAE (Multimodal Variational Autoencoder)	Texto da postagem e também a imagem anexada	Notícias de sites como Weibo e Twitter	Sim	78,2% a 82,4% de acordo com as bases

**Tabela 6. Análise dos artigos**

## REFERÊNCIAS

Aldwairi, Monther; Alwahedi, Ali; **Detecting Fake News in Social Media Networks**; 2018;

Gereme, Fantahun Bogale; Zhu, William; **Early Detection of Fake News "Before It Flies High"**; 2019;

Jang, S Mo; Geng, Tieming; Li, Jo-Yun Queenie; Xia, Ruofan; Huang, Chin-Tser; Kim, Hwalbin; Tang, Jijun; **A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis**; 2019;

Khattar, Dhruv; Goud, Jaipal Singh; Gupta, Manish; Varma, Vasudeva; **MVAE: Multimodal Variational Autoencoder for Fake News Detection**; 2019;

Ko, Hoon; Hong, Jong Youl; Kim, Sangheon; Mesicek, Libor; Na, In Seop; **Human-machine interaction: A case study on fake news detection using a backtracking based on a cognitive system**; 2019;

Ozbay, Feyza Altunbey; Alatas, Bilal; **Fake news detection within online social media using supervised artificial intelligence algorithms**; 2019;

Qian, Feng; Gong, Chengyue; Sharma, Karishma; Liu, Yan; **Neural User Response Generator: Fake News Detection with Collective User Intelligence**; 2018;

Tschiatschek, Sebastia; Singla, Adis; Gomez Rodriguez, Manuel; Merchant, Arpit; Krause, Andreas; **Fake News Detection in Social Networks via Crowd Signals**; 2018;

Vedova, Marco Della; Tacchini, Eugenio; Moret, Stefano; Ballarin, Gabriele; DiPierro, Massimo; de Alfaro, Luca; **Automatic Online Fake News Detection Combining Content and Social Signals**; 2018;

Zhang, Chaowei; Gupta, Ashish; Kauten, Christian; Deokar, Amit V; Qin, Xiao; **Detecting fake news for reducing misinformation risks using analytics approaches**; 2019;