

# Ciência da Computação

## O USO DE APRENDIZAGEM DE MÁQUINA PARA IDENTIFICAÇÃO E CLASSIFICAÇÃO DE FAKE NEWS NO TWITTER REFERENTES A ELEIÇÃO PRESIDENCIAL DE 2018

**Autor: Ithalo Henrique de Sousa Leal**

**Orientador: Glauber Luiz da Silva Costa**

**Caratinga, 2018**

# Agenda

- ▶ Introdução
- ▶ Referencial Teórico
- ▶ Objetivo Geral
- ▶ Metodologia
- ▶ Análise dos Resultados
- ▶ Conclusão
- ▶ Trabalhos Futuros
- ▶ Principais Referências

# Introdução

- ▶ Atualmente com o grande volume de dados, que são gerados todos os dias ao redor do mundo, surgiu a necessidade de exploração desses dados, com isso o mercado na área de mineração de dados vem crescendo a cada ano.
- ▶ A sociedade tem procurado maneiras práticas e rápidas para se comunicar e realizar suas tarefas cotidianas, com isso o número de internautas a cada ano tem crescido.
- ▶ As mídias sociais permitem aos usuários se interagirem, contudo, muitas informações que são disseminadas nas mídias sociais não são verdadeiras.
- ▶ O tema foi escolhido, devido à eleição presidencial que ocorreu neste ano de 2018, e as *fake news* foram um dos assuntos mais tratados durante toda a eleição, sendo uma das principais preocupações.

# Referencial teórico

## ▶ ***Fake News***

- ▶ O termo *fake news* traduzido para o português significa notícias falsas. Esse termo tem sido muito falado atualmente, inclusive a imprensa tem feito várias matérias para relatar sobre o tema. A criação de *fake news* tem por finalidade denegrir uma pessoa ou entidade. Geralmente são criadas matérias absurdas para atrair acesso. Segundo um estudo realizado Aral, Roy e Vosoughi (2018) notícias falsas se espalham de forma muito mais rápida no Twitter, sendo 70% mais propensas a serem compartilhadas do que as notícias verdadeiras.

## ▶ **KDD (*Knowledge Discovery in Databases*)**

- ▶ O KDD tem como objetivo a descoberta de conhecimento válido e relevante de forma automatizada em uma base de dados.

## ▶ **Mineração de Dados**

- ▶ A mineração de dados, ou no inglês *data mining*, tem sido utilizada em diversas áreas, como nas áreas educacionais, financeira e medicina, por sua capacidade em extrair informação que antes de sua aplicação não era possível.

# Referencial teórico

## ▶ **Mineração de textos**

- ▶ A mineração de texto pode ser considerada, uma busca de informações relevantes a partir de um grande volume de textos, escrito na linguagem natural. Tanto na mineração de dados quanto na mineração de textos são aplicados algoritmos de aprendizagem de máquina, visando o conhecimento (VIANA, 2014).

## ▶ **Aprendizagem de Máquina**

- ▶ De acordo com Goldschmidt (2010) o aprendizado de máquina é uma área da inteligência artificial voltada para o desenvolvimento de algoritmos e técnicas computacionais que possibilitam que os computadores sejam capazes de aprenderem.

# Referencial teórico

## ▶ WEKA

- ▶ WEKA é uma coleção de algoritmos de aprendizado de máquina para tarefa de mineração de dados, contendo inúmeras ferramentas para preparação, classificação, regressão e agrupamento de dados, a ferramenta começou a ser escrita em 1993 na Universidade de Waikato, Nova Zelândia.

## ▶ Twitter

- ▶ França e Oliveira (2014) define o Twitter como um microblog onde permitem pessoas e instituições a divulgar qualquer tipo de informação ou opinião de forma quase instantânea, para todos que estão conectados a sua rede. Cada publicação é limitada um número de caracteres, obrigando os usuários a expressar sua opinião ou dar uma informação de forma rápida e objetiva.

# Objetivo geral

- ▶ O objetivo desse trabalho foi o desenvolvimento de um sistema protótipo capaz de classificar e Identificar *fake news* referentes à eleição presidencial no Brasil do ano de 2018.

# Metodologia

- ▶ Base de dados para coleta dos *tweets*
- ▶ Geração do arquivo .ARFF
- ▶ Pré-processamento dos dados
  - ▶ *StringToWordVector*
  - ▶ *Reorder*
  - ▶ *Normalização*



# Metodologia

- ▶ Implementação dos algoritmos classificadores
  - ▶ *NaiveBayes*
  - ▶ MultilayerPerceptron
  - ▶ J48
  - ▶ Jrip
  - ▶ iBK
- ▶ Treinamentos dos algoritmos

# Análise de resultados

Tabela 2 - Quantidade de registros classificados em cada classe

	<b><i>NaiveBayes</i></b>	<b>J48</b>	<b>JRip</b>	<b>IBK</b>
<b>opinião</b>	1.342.164	1.390.367	1.696.925	1.348.893
<b>fato</b>	245.825	214.722	48.698	255.551
<b>fake</b>	192.832	175.732	35.198	176.377

Fonte: O autor (2018)

# Análise de resultados

Tabela 3 - Percentual de registros classificados em cada classe

	<b><i>NaiveBayes</i></b>	<b>J48</b>	<b>JRip</b>	<b>IBK</b>
<b>opinião</b>	75,36771%	78,07449%	95,28891%	75,74557%
<b>fato</b>	13,80403%	12,05747%	2,73458%	14,35018%
<b>fake</b>	10,82826%	9,86803%	1,97650%	9,90425%

Fonte: O autor (2018)

# Análise de resultados: *K-fold Cross Validation*

Tabela 4 - Testes percentual de acertos alterando o *seed*

	<b>NAIVEBAYES</b>	<b>J48</b>	<b>JRIP</b>	<b>IBK</b>
<b>Seed 1</b>	91,20491%	91,59836%	91,68032%	73,71311%
<b>Seed 2</b>	91,20491%	91,72131%	91,67213%	73,72131%
<b>Seed 3</b>	91,21311%	91,68032%	91,89344%	73,93442%
<b>Seed 4</b>	91,22950%	91,56557%	91,93442%	73,58196%
<b>Seed 5</b>	91,21311%	91,37704%	91,86065%	73,86065%
<b>Seed 6</b>	91,19672%	91,34426%	91,87704%	74,06557%
<b>Seed 7</b>	91,21311%	91,63934%	91,71311%	73,85245%
<b>Seed 8</b>	91,21311%	91,45901%	91,68852%	73,50000%
<b>Seed 9</b>	91,22131%	91,23770%	91,89344%	73,89344%
<b>Seed 10</b>	91,21311%	91,85245%	91,80327%	73,96721%
<b>MÉDIA:</b>	<b>91,21229%</b>	<b>91,54754%</b>	<b>91,80163%</b>	<b>73,80901%</b>

Fonte: O autor (2018)

# Análise de resultados: Matriz de confusão

Tabela 5 - Melhor desempenho algoritmo *NaiveBayes* com *seed* 4

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	4.159	0	29
<b>fato</b>	458	3.512	90
<b>fake</b>	416	77	3.459

Fonte: O autor (2018)

Tabela 6 - Pior desempenho algoritmo *NaiveBayes* com *seed* 6

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	4.158	0	30
<b>fato</b>	458	3.511	91
<b>fake</b>	415	80	3.457

Fonte: O autor (2018)

# Análise de resultados: Matriz de confusão

Tabela 7 - Melhor desempenho algoritmo J48 com *seed* 10

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	3.869	150	169
<b>fato</b>	238	3.754	68
<b>fake</b>	295	74	3.583

Fonte: O autor (2018)

Tabela 8 - Pior desempenho algoritmo J48 com *seed* 9

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	3.845	170	173
<b>fato</b>	272	3.724	64
<b>fake</b>	321	69	3.562

# Análise de resultados: Matriz de confusão

Tabela 9 - Melhor desempenho algoritmo JRip com seed 4

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	4.090	25	73
<b>fato</b>	419	3.586	55
<b>fake</b>	384	28	3.540

Fonte: O autor (2018)

Tabela 10 - Pior desempenho algoritmo JRip com seed 2

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	4.089	32	67
<b>fato</b>	437	3.569	54
<b>fake</b>	393	33	3.526

Fonte: O autor (2018)

# Análise de resultados: Matriz de confusão

Tabela 11 - Melhor desempenho algoritmo IBK com *seed* 6

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	3.981	126	81
<b>fato</b>	545	2.689	826
<b>fake</b>	1.518	68	2.366

Fonte: O autor (2018)

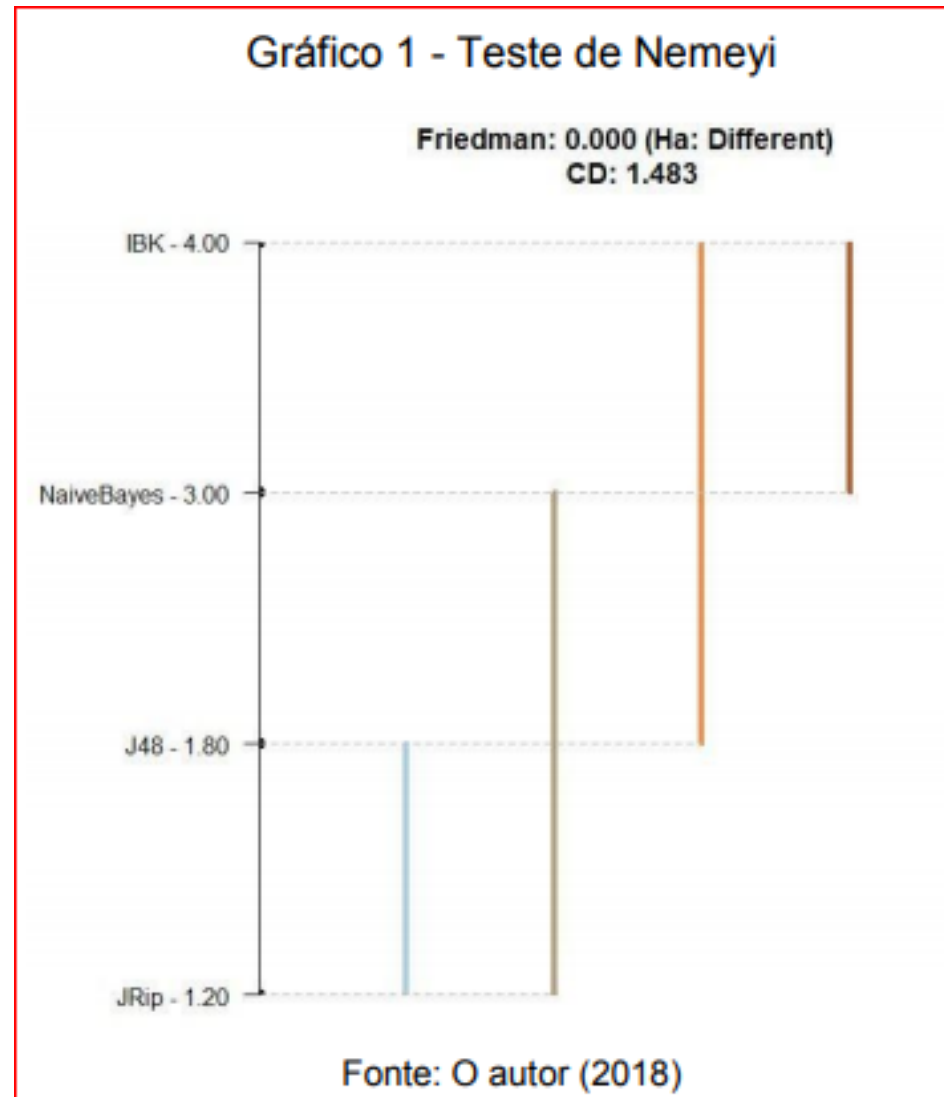
Tabela 12 - Pior desempenho algoritmo IBK com *seed* 8

	<b>opinião</b>	<b>fato</b>	<b>fake</b>
<b>opinião</b>	3.994	120	74
<b>fato</b>	545	2.660	855
<b>fake</b>	1.567	72	2.313

Fonte: O autor (2018)



# Análise de resultados: Teste de Friedman e Nemenyi



# Conclusão

- ▶ Através do desenvolvimento desse trabalho, foi possível validar a utilização de aprendizagem de máquina para o combate as *fake news*, haja visto que a classe *fake* obteve ótima precisão, sendo superior a 90% em três dos quatro algoritmos utilizados. Se analisarmos o algoritmo na totalidade, ainda sim obtemos uma acurácia acima dos 90% em três algoritmos.
- ▶ Efetuando um levantamento quantitativo dos registros classificados como *fake news*, foi possível notar-se que os três candidatos mais citados foram Cabo Daciolo sendo citado em 68.330 (sessenta e oito mil e trezentos e trinta) registros, seguido do candidato Fernando Haddad com 57.145 (cinquenta e sete mil e cento e quarenta e cinco) e por fim o candidato Henrique Meirelles, sendo citado em 47.795 (quarenta e sete mil e setecentos e noventa e cinco) registros.

# Conclusão

- ▶ Foi fundamental para os resultados obtidos nesse trabalho, a utilização do corpus disponibilizado por Monteiro et al. (2018), pois, os registros classificados manualmente estavam desbalanceados para a classe opinião afetando o desempenho dos algoritmos nas classes *fake* e fato.
- ▶ As *fake news* apesar de serem um grande problema a ser combatido, nota-se pouco material científico para resolução de forma automatizada. Pode-se concluir que o objetivo principal do trabalho em classificação e identificação de *fake news* obteve 60 sucesso. Os resultados obtidos foram além da expectativa do autor, haja visto que classificação de textos em português é uma atividade complexa.

# Trabalhos futuros

- ▶ Classificar mais registros de forma manual e efetuar novo treinamento e teste dos algoritmos para os reavaliar.
- ▶ Analisar a propagação de uma *fake news* identificada.
- ▶ Analisar uma influência nos resultados da eleição.
- ▶ Desenvolvimento de aplicativos utilizando o modelo criado para ajuda aos usuários na detecção de *fake news*.
- ▶ Testar outros algoritmos.
- ▶ Utilização de outras bases de dados para coleta de registros.

# Principais referências

- ARANHA, Christian Nunes. **Uma Abordagem de Pré-processamento Automático para Mineração de Textos em Português: Sob o Enfoque da Inteligência Computacional**. Rio de Janeiro, 2007 Tese (Engenharia Elétrica) - PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO, 2007.
- FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. **From Data Mining to Knowledge discovery**: American Association for Artificial Intelligence. 1996.
- GOLDSCHMIDT, Ronaldo Ribeiro. **Uma introdução à Inteligência Computacional: Fundamentos, Ferramentas e Aplicações**. 1. ed. Rio de Janeiro: Instituto Superior de Tecnologia do Rio de Janeiro, 2010.
- GOLDSCHMIDT, Ronaldo Ribeiro; PASSOS, Ronaldo; BEZERRA, Emmanuel. **Data Mining: Conceitos, técnicas, algoritmos, orientações e aplicações**. 2. ed. Elsevier Editora Ltda, 2015.
- MONTEIRO, R.A. et al. Contributions to the Study of Fake News in Portuguese: New Corpus and Automatic Detection Results. In: COMPUTATIONAL Processing of the Portuguese Language. Springer International Publishing, 2018, p. 324-334

Obrigado pela atenção!