

**Ernane Rosa Martins  
(Organizador)**

# **A Abrangência da Ciência da Computação na Atualidade**

**Ernane Rosa Martins**

(Organizador)

# A Abrangência da Ciência da Computação na Atualidade

Atena Editora  
2019

2019 by Atena Editora  
Copyright © Atena Editora  
Copyright do Texto © 2019 Os Autores  
Copyright da Edição © 2019 Atena Editora  
Editora Executiva: Prof<sup>a</sup> Dr<sup>a</sup> Antonella Carvalho de Oliveira  
Diagramação: Karine de Lima  
Edição de Arte: Lorena Prestes  
Revisão: Os Autores

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

### **Conselho Editorial**

#### **Ciências Humanas e Sociais Aplicadas**

Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas  
Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná  
Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília  
Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa  
Prof<sup>a</sup> Dr<sup>a</sup> Cristina Gaio – Universidade de Lisboa  
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia  
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná  
Prof<sup>a</sup> Dr<sup>a</sup> Ivone Goulart Lopes – Istituto Internazionele delle Figlie de Maria Ausiliatrice  
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense  
Prof<sup>a</sup> Dr<sup>a</sup> Lina Maria Gonçalves – Universidade Federal do Tocantins  
Prof<sup>a</sup> Dr<sup>a</sup> Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Prof<sup>a</sup> Dr<sup>a</sup> Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa  
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará  
Prof<sup>a</sup> Dr<sup>a</sup> Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

#### **Ciências Agrárias e Multidisciplinar**

Prof. Dr. Alan Mario Zuffo – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano  
Prof<sup>a</sup> Dr<sup>a</sup> Daiane Garabeli Trojan – Universidade Norte do Paraná  
Prof. Dr. Darllan Collins da Cunha e Silva – Universidade Estadual Paulista  
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul  
Prof<sup>a</sup> Dr<sup>a</sup> Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia  
Prof. Dr. Jorge González Aguilera – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará  
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

#### **Ciências Biológicas e da Saúde**

Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás  
Prof.<sup>a</sup> Dr.<sup>a</sup> Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina  
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria  
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará

Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão  
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

### **Ciências Exatas e da Terra e Engenharias**

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto  
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná  
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

### **Conselho Técnico Científico**

Prof. Msc. Abrãao Carvalho Nogueira – Universidade Federal do Espírito Santo  
Prof. Dr. Adaylson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba  
Prof. Msc. André Flávio Gonçalves Silva – Universidade Federal do Maranhão  
Prof.ª Drª Andreza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico  
Prof. Msc. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro  
Prof. Msc. Daniel da Silva Miranda – Universidade Federal do Pará  
Prof. Msc. Eliel Constantino da Silva – Universidade Estadual Paulista  
Prof.ª Msc. Jaqueline Oliveira Rezende – Universidade Federal de Uberlândia  
Prof. Msc. Leonardo Tullio – Universidade Estadual de Ponta Grossa  
Prof.ª Msc. Renata Luciane Polsaque Young Blood – UniSecal  
Prof. Dr. Welleson Feitosa Gazel – Universidade Paulista

<b>Dados Internacionais de Catalogação na Publicação (CIP) (eDOC BRASIL, Belo Horizonte/MG)</b>	
A161	A abrangência da ciência da computação na atualidade [recurso eletrônico] / Organizador Ernane Rosa Martins. – Ponta Grossa, PR: Atena Editora, 2019.  Formato: PDF Requisitos de sistema: Adobe Acrobat Reader Modo de acesso: World Wide Web Inclui bibliografia ISBN 978-85-7247-488-7 DOI 10.22533/at.ed.887190908  1. Computação – Pesquisa – Brasil. I. Martins, Ernane Rosa. CDD 004
<b>Elaborado por Maurício Amormino Júnior – CRB6/2422</b>	

Atena Editora  
Ponta Grossa – Paraná - Brasil  
[www.atenaeditora.com.br](http://www.atenaeditora.com.br)  
contato@atenaeditora.com.br

## APRESENTAÇÃO

A área da Ciência da Computação apresenta atualmente uma constante ascensão, seus profissionais estão sendo cada vez mais valorizados e requisitados pelas empresas, tornando-a mais importante, prestigiada e reconhecida. As empresas de todos os portes e setores necessitam de profissionais qualificados desta área, que apresentem potencial para promover inovação, desenvolvimento e eficiência.

A Ciência da Computação é uma área com amplas possibilidades de atuação, como por exemplo: a elaboração de programas e softwares, o gerenciamento de informações, a atuação acadêmica, a programação de aplicativos mobile ou ainda de forma autônoma. A abrangência da Ciência da Computação exige de seus profissionais conhecimentos diversos, tais como: novos idiomas, pensamento criativo, capacidade de comunicação e de negociação, além da necessidade de uma constante atualização de seus conhecimentos.

Dentro deste contexto, este livro aborda diversos assuntos importantes para os profissionais e estudantes desta área, tais como: API de localização da google, identificação de etiquetas RFID, ferramentas para recuperação de dados, ensino de computação, realidade virtual, interação humano computador, gestão do conhecimento, computação vestível, gerência de projetos, big data, mineração de dados, Internet das coisas, monitoramento do consumo de dados na Internet, pensamento computacional, análise de sentimentos, filtros ópticos, rede óptica elástica translúcida, algoritmo de roteamento, algoritmo de atribuição espectral, algoritmo de utilização de regeneradores e algoritmo genético.

Assim, certamente que os trabalhos apresentados nesta obra exemplificam um pouco a abrangência da área de Ciência da Computação na atualidade, permitindo aos leitores analisar e discutir os relevantes assuntos abordados. A cada autor, nossos agradecimentos por contribuir com esta obra, e aos leitores, desejo uma excelente leitura, repleta de boas reflexões.

Ernane Rosa Martins

## SUMÁRIO

<b>CAPÍTULO 1</b> .....	<b>1</b>
UMA ABORDAGEM SOBRE SISTEMA DE LOCALIZAÇÃO MOBILE	
Paulo Roberto Barbosa	
<b>DOI 10.22533/at.ed.8871909081</b>	
<b>CAPÍTULO 2</b> .....	<b>6</b>
UMA ABORDAGEM BIDINÂMICA PARA A IDENTIFICAÇÃO DE ETIQUETAS RFID	
Shalton Viana dos Santos	
Paulo André da S. Gonçalves	
<b>DOI 10.22533/at.ed.8871909082</b>	
<b>CAPÍTULO 3</b> .....	<b>23</b>
TESTE DE FERRAMENTAS DE RECUPERAÇÃO DE IMAGENS PARA SISTEMAS DE ARQUIVOS EXT3 E EXT4	
Diego Vinícius Natividade	
<b>DOI 10.22533/at.ed.8871909083</b>	
<b>CAPÍTULO 4</b> .....	<b>34</b>
REDIMENSIONAMENTO DO ENSINO DA COMPUTAÇÃO NA EDUCAÇÃO BÁSICA: O PENSAMENTO COMPUTACIONAL, O UNIVERSO E A CULTURA DIGITAL	
Melquisedec Sampaio Leite	
Sônia Regina Fortes da Silva	
<b>DOI 10.22533/at.ed.8871909084</b>	
<b>CAPÍTULO 5</b> .....	<b>47</b>
REALIDADE VIRTUAL, UTILIZANDO DAS MELHORES PRÁTICAS DA INTERAÇÃO HUMANO COMPUTADOR	
Bruno Moreira Batista	
Guiliano Rangel Alves	
Hellen Corrêa da Silva	
Rhogério Correia de Souza Araújo	
<b>DOI 10.22533/at.ed.8871909085</b>	
<b>CAPÍTULO 6</b> .....	<b>52</b>
ORGANIZAÇÃO DO CONHECIMENTO PARA A MEMÓRIA EMPRESARIAL: UM RELATO TÉCNICO SOBRE A EXPERIÊNCIA DO SEBRAE/RJ	
Leandro Pacheco de Melo	
<b>DOI 10.22533/at.ed.8871909086</b>	
<b>CAPÍTULO 7</b> .....	<b>65</b>
GERÊNCIA DE PROJETOS EM COMPUTAÇÃO VESTÍVEL: DIRETRIZES PARA O DESENVOLVIMENTO DE PRODUTOS VESTÍVEIS INTELIGENTES	
Renan Gomes Barreto	
Lucas Oliveira Costa Aversari	
Renata Gomes Barreto	
Gabriela Ferreira Marinho Barreto	
<b>DOI 10.22533/at.ed.8871909087</b>	

<b>CAPÍTULO 8</b> .....	<b>76</b>
EXPLORING <i>BIG DATA</i> CONTENT AND INFORMATION METRICS: INTERSECTIONS AND ANALYSIS TO SUPPORT DECISION-MAKING	
Rafael Barcellos Gomes Vânia Lisboa da Silveira Guedes	
<b>DOI 10.22533/at.ed.8871909088</b>	
<b>CAPÍTULO 9</b> .....	<b>92</b>
DEMOCHAIN - FRAMEWORK DESTINADO A CRIAÇÃO DE REDES BLOCKCHAIN HÍBRIDAS PARA DISPOSITIVOS IOT	
Lorenzo W. Freitas Carlos Oberdan Rolim	
<b>DOI 10.22533/at.ed.8871909089</b>	
<b>CAPÍTULO 10</b> .....	<b>107</b>
CONSUMO DO TRÁFEGO DE DADOS EM APLICAÇÕES DE VÍDEO SOB DEMANDA- YOUTUBE E NETFLIX	
Patricia Emilly Nóbrega da Silva Éwerton Rômulo Silva Castro	
<b>DOI 10.22533/at.ed.88719090810</b>	
<b>CAPÍTULO 11</b> .....	<b>112</b>
COMPUTAÇÃO NA ESCOLA: ABORDAGEM DESPLUGADA NA EDUCAÇÃO BÁSICA	
Christian Puhmann Brackmann Marcos Román-González Rafael Marimon Boucinha Dante Augusto Couto Barone Ana Casali Flávia Pereira da Silva	
<b>DOI 10.22533/at.ed.88719090811</b>	
<b>CAPÍTULO 12</b> .....	<b>128</b>
COLETA DE DADOS E ANÁLISE DE SENTIMENTOS NAS REDE SOCIAIS ON LINE	
Maurilio Alves Martins da Costa Bruna Emidia de Assis Almeida Fraga	
<b>DOI 10.22533/at.ed.88719090812</b>	
<b>CAPÍTULO 13</b> .....	<b>137</b>
ANÁLISE DO IMPACTO DO CASCATEAMENTO DE FILTROS ÓPTICOS EM UM CENÁRIO DE REDES ÓPTICAS ELÁSTICAS	
Gabriela Sobreira Dias de Carvalho William Silva dos Santos Lucas Oliveira de Figueiredo Helder Alves Pereira	
<b>DOI 10.22533/at.ed.88719090813</b>	

<b>CAPÍTULO 14</b> .....	<b>143</b>
ANÁLISE DE REDE ÓPTICA ELÁSTICA TRANSLÚCIDA CONSIDERANDO DIFERENTES ALGORITMOS DE ROTEAMENTO	
Arthur Hendricks Mendes de Oliveira	
William Silva dos Santos	
Helder Alves Pereira	
Raul Camelo de Andrade Almeida Júnior	
<b>DOI 10.22533/at.ed.88719090814</b>	
<b>CAPÍTULO 15</b> .....	<b>149</b>
ANÁLISE DE REDE ÓPTICA ELÁSTICA TRANSLÚCIDA CONSIDERANDO ALGORITMOS DE ATRIBUIÇÃO ESPECTRAL	
Arthur Hendricks Mendes de Oliveira	
William Silva dos Santos	
Helder Alves Pereira	
Raul Camelo de Andrade Almeida Júnior	
<b>DOI 10.22533/at.ed.88719090815</b>	
<b>CAPÍTULO 16</b> .....	<b>155</b>
A NEW MULTI OBJECTIVE APPROACH FOR OPTIMIZING P-MEDIAN MODELING IN SCHOOL ALLOCATION USING GENETIC ALGORITHM	
Clahildek Matos Xavier	
Marly Guimarães Fernandes Costa	
Cícero Ferreira Fernandes Costa Filho	
<b>DOI 10.22533/at.ed.88719090816</b>	
<b>SOBRE O ORGANIZADOR</b> .....	<b>168</b>
<b>ÍNDICE REMISSIVO</b> .....	<b>169</b>



## COLETA DE DADOS E ANÁLISE DE SENTIMENTOS NAS REDE SOCIAIS ON LINE

**Maurilio Alves Martins da Costa**  
**Bruna Emidia de Assis Fraga**

**RESUMO:** As redes sociais on line (RSO) são utilizadas por milhões de usuários para compartilharem pensamentos, acontecimentos e opiniões sobre fatos e produtos. Este trabalho se objetiva a utilizar a técnica de análise de sentimentos (AS) para extrair o sentimento contido nas mensagens publicadas pelos usuários de uma rede social. Utilizou-se a abordagem léxica com saída polarizada [Benevenuto, Ribeiro e Araujo 2015] aplicada à RSO Twitter. Os resultados das coletas refletiram de modo satisfatório ao sentimento que correspondiam ao evento estudado, confirmando a polaridade existente bem como os destaques ocorridos nos debates.

**ABSTRACT:** Online social networks (OSN) are used by millions of persons to share thoughts, events, and opinions about facts and products. This work aims to use the technique of sentiment analysis (SA) to extract the sentiment contained in messages published by social network's users. We used the lexical approach with polarized output [Benevenuto, Ribeiro and Araujo 2015] applied to Twitter. The results reflected satisfactorily the feeling that corresponded to the event studied, confirming

the existing polarity as well as the highlights of the presidential debates..

### 1 | INTRODUÇÃO

As redes sociais online (RSO) são um tipo específico de mídia social. De acordo com Kaplan e Haenlein (2010) as redes sociais online possuem a funcionalidade de compartilhar conteúdos e têm seu foco voltado para o relacionamento e o engajamento entre os usuários de forma que seja estabelecida uma interação virtual entre eles.

As redes sociais online são em sua maioria gratuitas e alcançam em torno de 3 bilhões de pessoas em todo o mundo. Elas foram desenvolvidas para que as pessoas criem perfis, estabeleçam uma rede de amigos e publiquem conteúdos. Estes conteúdos podem ser de diversos formatos como links, fotos e vídeos mas principalmente mensagens textuais.

O conteúdo produzido por um usuário é conhecido como *user-generated content* (UGC) e dada a riqueza de informações contidas nas publicações geradas sobre variados temas e assuntos Balazs e Velásquez (2016) afirmam que é humanamente impossível compreender toda essa informação em um período de tempo satisfatório, sendo necessários mecanismos para extrair e analisar esses conteúdos,

surgindo, então, o campo de pesquisa da análise de sentimento (AS) ou mineração de opinião (MO).

Segundo Liu (2012), a análise de sentimento permite que se tenha uma percepção geral da opinião dos usuários da rede em um dado momento sobre um assunto, ou ainda a captura da reação do público, seja ela positiva ou negativa, sobre algum tema. De acordo com os estudos realizados por Medhat, Hassan e Korashy (2014) a análise de sentimentos pode ser realizada a partir de duas abordagens principais: (i) uma baseada em aprendizado de máquina e outra baseada em análise léxica. O aprendizado de máquina pode ser definido como um conjunto de algoritmos capazes de reconhecer padrões a partir de um conjunto de dados e fazer uma previsão futura. (ii) A outra abordagem para a análise de sentimentos é a análise léxica, que segundo Benevenuto, Ribeiro e Araújo (2015), relaciona valores quantitativos ou qualitativos às palavras para atribuir-lhes sentimento. Assim, é formado um conjunto de palavras opinativas que compõe um dicionário léxico. Nesta linha de abordagem, podem ser utilizados também padrões sintáticos para reconhecer novas palavras além daquelas que já existem no conjunto [Medhat, Hassan e Korashy 2014]. Segundo esses autores, a abordagem léxica é vantajosa porque não necessita de classificação prévia de sentenças e nem de base de treino como ocorre nos métodos de aprendizagem de máquina, assim, ela fica livre de contexto por justamente não sofrer especialização durante o treinamento

Da análise de sentimentos obtêm-se como resultado uma classificação do sentimento que pode envolver três aspectos: polaridade, emoção e força do sentimento [Benevenuto, Ribeiro e Araújo 2015]. Segundo os autores, a saída comum dos métodos é a polarizada. A polaridade indica uma orientação que pode ser positiva, negativa ou neutra, assim, o texto analisado terá uma saída que pode ser binária (positiva ou negativa) ou terciária (positiva, negativa e neutra).

O objetivo deste estudo é descrever uma metodologia de coleta e extração do sentimento exposto nas mensagens textuais publicadas pelos usuários de uma rede social, utilizando um método de análise de sentimentos com a abordagem de análise léxica com saída polarizada.

## 2 | ESTADO DA ARTE

A ascensão dos problemas de análise de sentimentos levou à publicação de um grande volume de pesquisas. As áreas de destaque destas publicações foram: a de saúde [Chew e Eysenbach 2010] e [Vydiswaran, Zhai e Roth 2011]; área empresarial, com o monitoramento do sentimento do público em relação às marcas e produtos [Evangelista e Padilha 2013] e [Ghiassi, Skinner e Zimbra 2013], social [Nascimento et al. 2012] e [Malheiros e Lima 2013] e política [Stieglitz e Dang-Xuan, 2012] e [Malini, Ciarelli e Medeiros 2017].

Em Cavalin et al. (2014), foi desenvolvido um sistema de tempo real capaz de

analisar o sentimento das mensagens dos usuários do Twitter. Ao longo da pesquisa foram analisadas as mensagens durante os jogos da Copa das Conferências de 2013. Quanto à análise de sentimentos a abordagem de aprendizado de máquina foi utilizada e o classificador probabilístico de Naive Bayes polarizou as mensagens em sentimentos positivo, negativo e neutro.

Souza (2015) criou um método para identificar usuários influentes da rede social Twitter e classificar automaticamente a categoria na qual o usuário é influente, baseando-se nos conteúdos postados. Os conteúdos poderiam pertencer a três categorias: política, esporte ou outros assuntos. Para a classificação das postagens, utilizou-se técnicas de mineração de texto e o método de aprendizado supervisionado Naive Bayes.

Wang et al. (2012) desenvolveu um sistema em tempo real que monitorou as eleições presidenciais americanas de 2012 no Twitter. Em seu estudo foram considerados quatro tipos de sentimentos: negativo, positivo, neutro e incerto. Para formar as bases de treino e teste do modelo Naive Bayes (NB) unigrama, uma série de mensagens foram classificadas por voluntários.

### 3 | FUNDAMENTOS TEÓRICOS

#### 3.1 Redes Sociais Online (RSO)

As RSOs podem ser expressas como serviços baseados na web no qual as pessoas podem criar perfis em um sistema, podendo ser eles públicos ou não [Boyd e Ellison 2007]. Ainda de acordo com o autor, através desse serviço os usuários se apresentam, conectam-se uns aos outros visualizando e explorando tanto à sua lista de conexões como as de outras pessoas. O principal objetivo das plataformas sociais é formar e tornar visíveis as conexões entre pessoas.

Apesar de possuírem um objetivo central em comum, cada RSO disponibiliza recursos e funcionalidades de acordo com o seu propósito. Por exemplo, o Facebook<sup>1</sup> é têm como propósito a criação de relacionamentos de amizade. Já o Pinterest<sup>2</sup> e o Instagram<sup>3</sup> incentivam seus usuários à compartilharem imagens. Com foco nos negócios a rede LinkedIn<sup>4</sup> permite que seus participantes se apresentem de forma profissional revelando aptdões e habilidades, compartilhem conteúdos relacionados à temas profissionais relacionados ao trabalho e favorece o networking. Em uma outra frente, têm-se o Twitter<sup>5</sup>, uma rede que foca em conteúdos e notícias rápidas.

As redes Facebook e Twitter receberão maior destaque ao longo deste estudo por se tratar de redes que possuem características relacionadas à mensagens textuais e público mais coerentes com os objetivos deste estudo.

- 1 <https://www.facebook.com>
- 2 <https://br.pinterest.com>
- 3 <https://www.instagram.com>
- 4 <https://br.linkedin.com>
- 5 <https://twitter.com>

O Facebook estabelece a conexão entre os usuários de forma bidirecional, assim um usuário solicita a amizade e outro deve confirmar. As amizades geralmente já existem no meio offline e posteriormente se tornam virtuais. No Twitter, ocorre a unidirecionalidade da conexão, em que só existe a opção seguir. Dessa forma, os usuários buscam seguir e receber atualizações somente daqueles que produzem conteúdos considerados para eles relevantes. Assim, o Facebook conecta pessoas, enquanto o Twitter conecta tópicos e idéias.

As duas redes possuem em comum recursos como hashtags, menções, tópicos e é possível pesquisar pessoas, negócios, tópicos e empresas. Ainda segundo o autor o Facebook ao disponibilizar mais recursos que o Twitter, se torna mais difícil de ser usado pelos usuários.

### 3.2 Extração de dados das Redes Sociais Online

As redes sociais em geral, disponibilizam interfaces para a interação ou extração de seus dados. Essas interfaces são conhecidas como Application Programming Interface (API) e para realizar a extração de dados é preciso conhecer suas especificidades.

A principal API da rede social Facebook é a Graph API, disponível em <https://developers.facebook.com/>. Segundo documentação disponível em <https://developers.facebook.com/docs/graph-api/>, todas as outras APIs existentes são expansões desta. A Graph API estabelece uma limitação de volume de requisições que podem ser feitas ao servidor dentro de um intervalo de tempo. Durante uma hora, uma aplicação pode realizar uma quantidade de chamadas igual a duzentas vezes o número de usuários que ela possui. Para realizar solicitações à Graph API são necessários tokens de acesso que fornecem autorização para tal. Uma outra API é a de Feed Público, que necessita de aprovação prévia do Facebook.

O Twitter fornece aos desenvolvedores uma grande variedade de APIs para a construção de aplicações ou soluções para a sua ferramenta. A API Standard, que será utilizada neste trabalho, é uma categoria de API gratuita, já as duas outras categorias, Premium e Enterprise, são pagas. A categoria Standard conta com dois módulos: (i) API REST que permite a realização de consultas no banco de dados do Twitter de dados publicados há no máximo sete dias e; (ii) API Streaming que permite aos desenvolvedores acesso às publicações em tempo real.

### 3.3 Mineração de Texto

A Mineração de textos (MT) é uma extensão da área de mineração de dados [Hearst 1999]. Esta subárea foi desenvolvida para que a partir de textos não estruturados ou semi-estruturados, fosse possível encontrar padrões relevantes e não triviais e ainda extrair novas informações [Kantardzic 2011]. Nesse processo, conteúdos textuais de diferentes fontes são extraídos automaticamente com o objetivo de revelar uma informação completamente nova.

A mineração de dados e a mineração de textos, se diferenciam na forma de reconhecer os padrões de seus objetos de análise. Para minerar uma forma textual, segundo Weiss et al. (2010), se faz necessária a transformação dos textos em uma representação numérica. Já na mineração de dados, as informações estão estruturadas dentro do banco de dados e são formadas por fatos, já que foram armazenadas automaticamente por sistemas [Hearst 2003].

Este trabalho utilizará a mineração de textos a fim de extrair os sentimentos dos documentos online produzidos por usuários das redes sociais. Para isto, foi escolhido o modelo de mineração textual proposto por [Aranha e Vellasco 2007]. Este modelo é formado por cinco etapas sequenciais. A primeira delas é a formação da base de documentos, cujo objetivo é a coleta de informações para a construção de um arcabouço de dados a serem explorados. Nesse trabalho a ênfase da coleta estará voltada para dados contidos na Internet.

A segunda etapa é o pré-processamento onde os dados são transformados para se adequarem a formatos passíveis de serem interpretados pelos algoritmos de extração automática de conhecimento. Esta etapa reduz o custo da etapa de mineração agindo na solução de problemas como dados ruidosos, faltantes e ou redundantes.

Posteriormente têm-se a fase da indexação que cria coleções de termos extraídos do conjunto de dados proporcionando otimização da velocidade e desempenho na busca realizada [Soares 2013].

A quarta fase é a da mineração que irá trabalhar em cima dos dados provenientes dos processos anteriores para a produção efetiva de novos conhecimentos úteis, aplicando sobre os dados, algoritmos pertencentes à diversas áreas do conhecimento.

A quinta fase é a da análise dos resultados obtidos, nessa fase ocorre a análise e interpretação dos resultados a partir dos dados obtidos, fazendo uma verificação para avaliar o classificador de acordo com algumas métricas que podem ser taxa de erro, tempo de CPU e complexidade do modelo ou ainda as métricas podem ser qualitativas, levando em consideração o conhecimento de especialistas do domínio.

#### **4 | METODOLOGIA**

O presente estudo é uma pesquisa caráter exploratório e descritivo. Cujo objetivo é conhecer o objeto de estudo de forma a compreender a sua contextualização e o seu significado. Este tipo de pesquisa ainda pode ser considerada uma “investigação em área onde há pouco conhecimento sistematizado, acumulado.” (Vergara 1990, p.4). Já a pesquisa descritiva, é definida por Vergara (1990) como sendo uma exposição de características de um fenômeno em que fatores e variáveis podem ser correlacionados.

Dentre as diversas redes sociais online disponíveis online, o Twitter foi escolhida como objeto de estudo por ter como principal objetivo a propagação de mensagens textuais em um espaço limitado, o que auxilia no processo de mineração do texto. Além do mais os perfis criados nessa rede social são definidas por padrão como públicas,

assim, todas as informações compartilhadas são visíveis à toda a rede. Por fim, o Twitter oferece uma interface mais acessível que as outras redes sociais online.

Como tema para estudo foi escolhido um assunto que tivesse o potencial de gerar discussões no Twitter, sendo escolhido assim a campanha presidencial de 2018. Para isto, foram escolhidos alguns eventos para o monitoramento das mensagens publicadas nas redes sociais, como mostra a Tabela 1.

<b>Evento</b>	<b>Período da Coleta (2018)</b>
Oficialização do Haddad como candidato	11 de Setembro
Debate no SBT (TV)	26 de Setembro
Debate na Record (TV)	30 de Setembro
Debate na Globo (TV)	04 de Outubro

Tabela 1 – Acontecimentos monitorados durante o período de eleições em 2018

Fonte: Autora

Esta pesquisa utilizou menções e hashtags como alvo dos algoritmos de busca por concentrarem as mensagens do tema em questão que foram: a oficialização da candidatura de Haddad e os debates nas emissoras SBT, Record e Globo. Ao final do período de coleta, a extração obteve como resultado quatro grandes arquivos com um total de 651.837 tweets. Além de informações referentes à própria mensagem como texto, menções e hashtags, é possível obter dados de usuário como: nome, quantidade de seguidores e seguidos, status de perfil, quantidade de curtidas bem como outras informações de perfil e configurações de conta. Os dados selecionados para serem minerados foram: o id do tweet – número que identifica a mensagem como única. o texto - a mensagem compartilhada na rede social. O nome dos usuários e das hashtags mencionados no tweet como também a data de publicação da mensagem.

Utilizou-se a ferramenta Sentiment140<sup>6</sup> para a classificação dos tweets conseguindo por consequência a polarização dos sentimentos. Porém, como a ferramenta opera na língua inglesa toda a base de dados necessitava tradução. Para tanto utilizou-se o Google Tradutor e por isso, foi necessário dividir a base de dados de cada evento em fragmentos menores para serem submetidos à tradução. Logo, reduziu-se a base de dados à um total de 20.000 tweets, sendo recolhidos 5.000 tweets para cada um dos eventos monitorados.

A partir de então aplicou-se o modelo de mineração textual proposto por Aranha e Vellasco (2007).

## 5 | RESULTADOS E DISCUSSÕES

A política é um tema que gera muitas discussões muitas das vezes acaloradas. As pessoas defendem os seus pontos de vista, levantam os prós e contras tanto das propostas como das qualidades de seus candidatos, elas fazem das palavras a sua

6 <http://www.sentiment140.com/>

maior arma para que a mensagem propagada se dissemine e consiga convencer os demais. Porém, o resultado das classificações mostrou que durante os debates e o evento de candidatura do Fernando Haddad, a maioria das mensagens publicadas eram fatos ou opiniões neutras, sem sentimentos explícitos.

A polaridade neutra prevaleceu com resultado acima dos 70% em ambas as classificações. Uma mensagem pode ser considerada neutra, devido à ausência de sentimentos expressos, ou também à presença equilibrada de sentimentos positivos e negativos. E em alguns casos, a polaridade neutra é atribuída a mensagens que não conseguiram ser classificadas nos outros sentimentos.

Outro fator que impacta na classificação dos sentimentos é a forma de escrita da internet. Ela dificulta ainda mais o processo, devido à linguagem predominantemente informal, uso de gírias, expressões populares e abreviações. Dessa forma, algumas palavras não foram traduzidas pelo tradutor e nem reconhecidas pelo classificador, interferindo no resultado.

Os candidatos mais citados durante o debate do SBT, foram Cabo Daciolo e Ciro Gomes. E os candidatos que mais receberam comentários positivos, foram respectivamente: Ciro Gomes, Haddad e Cabo Daciolo. Daciolo, foi citado precisamente 795 vezes e ganhou grande repercussão neste dia após uma de suas declarações no debate. Ciro Gomes também teve muitos tweets positivos (118 ao total).

O debate da Record dentre os demais debates foi o que apresentou o menor número de citações aos candidatos e o que apresentou a menor quantidade de sentimentos classificados. Os candidatos mais citados neste evento foram Ciro Gomes, Bolsonaro e Cabo Daciolo. Ciro Gomes, porém não recebeu nenhuma citação positiva neste debate só tweets negativos e neutros. Bolsonaro recebeu apenas tweets neutros e apesar de não ter comparecido ao debate foi o segundo mais comentado no Twitter.

O Debate da Globo apresentou os sentimentos explícitos de forma mais distribuída. Todos os candidatos que foram citados possuíram sentimentos positivos e negativos. Os candidatos mais citados neste debate foram Guilherme Boulos e Ciro Gomes. O candidato mais citado em tweets positivos foi novamente Ciro e o mais citado em tweets negativos foi o Fernando Haddad.

Quando se analisou o resultado do monitoramento dos debates percebeu-se que o Cabo Daciolo foi muito citado no debate do SBT e na Record, porém, o mesmo não foi convidado para o debate da Globo provocando uma queda drástica com o número de suas citações únicas nas mensagens. Entretanto, o mesmo não ocorreu com Bolsonaro que esteve ausente em todos os debates, porém, a quantidade de mensagens que o citavam foram sempre significativas quando comparadas aos candidatos presentes.

Através deste estudo foi possível se confirmar que o resultado obtido através das mensagens da rede social online, se referem aos sentimentos dos usuários no momento dos eventos, já que foi possível captar a reação instantânea do público às frases e discursos dos candidatos durante o debate.

## REFERENCES

- Aranha, C. N. e Vellasco, M. M. B. R..(2007). Uma abordagem de pré-processamento automático para mineração de textos em português: sob o enfoque da inteligência computacional. *Pontifícia Universidade Católica do Rio de Janeiro*, Rio de Janeiro, RJ.
- Balazs, J. A., e Velásquez, J. D. (2016). Opinion mining and information fusion: a survey. *Information Fusion*, 27, 95-110.
- Benevenuto, F., Ribeiro, F. e Araújo, M. (2015). Métodos para análise de sentimentos em mídias sociais. In *Brazilian Symposium on Multimedia and the Web (Webmedia)*, Manaus, Brasil.
- Boyd, D. M., e Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication*, 13(1), 210-230.
- Cavalin, P. R., Gatti, M. D. C., dos Santos, C. N., e Pinhanez, C. (2014). Real-time sentiment analysis in social media streams: The 2013 confederation cup case. *Proceedings of BRACIS/ENIAC*.
- Chew, C., e Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*, 5(11), e14118.
- Evangelista, T. R., e Padilha, T. P. P. (2013). Monitoramento de posts sobre empresas de ecommerce em redes sociais utilizando análise de sentimentos. In *Brazilian Workshop on Social Network Analysis and Mining (BraSNAM)*.
- Hearst, M. A. (1999, June). Untangling text data mining. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics* (pp. 3-10). Association for Computational Linguistics.
- Hearst, M. (2003). What is text mining. *SIMS*, UC Berkeley.
- Ghiassi, M., Skinner, J., e Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with applications*, 40(16), 6266-6282.
- Kantardzic, M. (2011). *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons.
- Kaplan, A. M. e Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. In *Business Horizons*, v. 53, n. 1, p. 59–68.
- Liu, B., e Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463). Springer, Boston, MA.
- Malheiros, Y. U. R. I., e Lima, G. E. O. R. G. E. (2013). Uma ferramenta para análise de sentimentos em redes sociais utilizando o sentinet. *IX Simpósio Brasileiro de Sistemas de Informação (SBSI2013)*.
- Malini, F., Ciarelli, P., e Medeiros, J. (2017). O sentimento político em redes sociais: big data, algoritmos e as emoções nos tweets sobre o impeachment de Dilma Rousseffl Political sentiment in social networks: big data, algorithms and emotions in tweets about the impeachment of Dilma Rouseff. *Liinc em Revista*, 13(2).
- Medhat, W., Hassan, A., e Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), 1093-1113.
- Nascimento, P., Aguas, R., De Lima, D., Kong, X., Osiek, B., Xexéo, G., e De Souza, J. (2012, July). Análise de sentimento de tweets com foco em notícias. In *Brazilian Workshop on Social Network Analysis and Mining*.



Soares, F. (2013). *Categorização Automática de Textos Baseada em Mineração de Textos* (Doctoral dissertation, PUC-Rio).

Souza, A. O. (2015) Identificação e classificação de usuários influentes no twitter por áreas de influência. TCC. Centro Federal de Educação Tecnológica de Minas Gerais.

Stieglitz, S., e Dang-Xuan, L. (2012, January). Political communication and influence through microblogging--An empirical analysis of sentiment in Twitter messages and retweet behavior. In *2012 45th Hawaii International Conference on System Sciences* (pp. 3500-3509). IEEE.

Vergara, S. C. (1990). *Tipos de pesquisa em administração*. Escola Brasileira de Administração Pública da FGV. Rio de Janeiro. RJ

Vydiswaran, V. G., Zhai, C., e Roth, D. (2011, August). Gauging the internet doctor: ranking medical claims based on community knowledge. In *Proceedings of the 2011 workshop on Data mining for medicine and healthcare* (pp. 42-51). ACM.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., e Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In *Proceedings of the ACL 2012 System Demonstrations* (pp. 115-120). Association for Computational Linguistics.

Weiss, S. M., Indurkha, N., Zhang, T., e Damerau, F. (2010). *Text mining: predictive methods for analyzing unstructured information*. Springer Science & Business Media.

## **SOBRE O ORGANIZADOR**

**Ernane Rosa Martins** - Doutorado em andamento em Ciência da Informação com ênfase em Sistemas, Tecnologias e Gestão da Informação, na Universidade Fernando Pessoa, em Porto/Portugal. Mestre em Engenharia de Produção e Sistemas, possui Pós-Graduação em Tecnologia em Gestão da Informação, Graduação em Ciência da Computação e Graduação em Sistemas de Informação. Professor de Informática no Instituto Federal de Educação, Ciência e Tecnologia de Goiás - IFG (Câmpus Luziânia) ministrando disciplinas nas áreas de Engenharia de Software, Desenvolvimento de Sistemas, Linguagens de Programação, Banco de Dados e Gestão em Tecnologia da Informação. Pesquisador do Núcleo de Inovação, Tecnologia e Educação (NITE), certificado pelo IFG no CNPq. ORCID: <https://orcid.org/0000-0002-1543-1108>

## ÍNDICE REMISSIVO

### B

Big data 76, 77

### C

Computação 2, 5, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 65, 67, 92, 112, 114, 115, 116, 117, 119, 120, 124, 127, 168, 169

Computação vestível 67

Comunicação 39, 42, 68, 75, 148, 154

Conhecimento 52, 53, 58, 59, 60, 61, 62, 70, 76

### D

Dispositivos 35

### E

Ensino 34, 35, 37, 40, 41, 42, 43, 45, 46, 107, 125, 127

### G

Gestão do conhecimento 63

### I

Informação 39, 52, 53, 56, 57, 58, 61, 63, 76, 89, 90, 91, 148, 154, 168

Internet 5, 7, 21, 22, 43, 57, 58, 92, 105, 106, 107, 112, 113, 115, 132

Internet das coisas 5

### M

Monitoramento 135

### O

Organização do conhecimento 54

### P

Programação 43, 168

### R

Recuperação de dados 24

Redes 21, 43, 130, 131, 137, 141, 148, 153, 154

### S

Sistemas de arquivos 24, 33

### T

Tecnologia 57, 60, 75, 112, 143, 148, 149, 154, 168

Testes 26, 27, 28, 29, 30, 31, 32, 122

Agência Brasileira do ISBN  
ISBN 978-85-7247-488-7

