

SISTEMA DE ANÁLISE DE SENTIMENTO EM COMENTÁRIOS RELACIONADOS A UNIVERSIDADES DE ENSINO SUPERIOR EM REDES SOCIAIS (VENERA)¹

Wender Lucas Souza²

José Corrêa Viana³

Fernando Corrêa de Mello Júnior⁴

Henaldo Barros Moraes⁵

RESUMO: Este artigo aborda o estudo e a aplicação de análise de sentimentos em cima de comentários de usuários de redes sociais referentes a instituições de ensino superior a fim de gerar indicadores de satisfação e dados volumétricos visando o apoio e suporte à criação de estratégias de divulgação e marketing. Assim, será apresentado o processo de criação de agentes coletores desses comentários na internet, de módulos de análise de sentimento desses comentários coletados e da interface final do usuário para visualização e manipulação desses dados analisados. Para o desenvolvimento do trabalho foram utilizadas técnicas de *webscraping* e *webcrawling* utilizando a linguagem *Python*, serviços de análise de sentimentos do *Microsoft Azure Cognitive Services* com o uso de *NodeJS* e os *frameworks React* e *VueJS* para o desenvolvimento da interface final.

PALAVRAS-CHAVE: Marketing Digital e Social; Análise de Sentimento; Análise de Indicadores.

ABSTRACT: This article will focus on the study and application of sentiment analysis based on comments from users of social networks referring to higher education institutions in order to generate satisfaction indicators and volumetric data aiming to support the creation of dissemination strategies and marketing. Thus, the process of creating agents collecting these comments on the internet, of sentiment analysis modules of these collected comments and of the final user interface for viewing and manipulating these analyzed data will be presented. For the development of the work, webscraping and webcrawling techniques were used using the Python language, sentiment analysis services from Microsoft Azure Cognitive Services using NodeJS and the React and VueJS frameworks for the development of the final interface.

KEYWORDS: Social and Digital Marketing; Sentiment Analysis; Indicator Analysis.

1 INTRODUÇÃO

A mídia sempre foi um ótimo modo de exposição e divulgação de imagem, seja ela de forma pessoal, empresarial ou outros afins. A televisão e o rádio tiveram papéis significativos nessa empreitada, pois após a chegada da internet, tornaram-se

¹ Artigo apresentado como avaliação final da disciplina de Estágio Supervisionado II, para obtenção de título de bacharel em Sistemas de Informação, pelo Centro Universitário de Patos de Minas – UNIPAM.

² Estudante de Graduação do Curso de Sistemas de Informação do UNIPAM. E-mail: wenderls@unipam.edu.br.

³ Professor orientador. E-mail: jcorrea@unipam.edu.br.

⁴ Professor orientador. E-mail: fernandocmj@unipam.edu.br.

⁵ Professor orientador. E-mail: henaldo@unipam.edu.br.

veículos de menor intensidade de compartilhamento de informação, de acordo com o crescimento de tecnologias mais flexíveis como celulares, tablets e computadores e aplicativos que fazem com que as pessoas resolvam tudo de maneira prática e simples, como o *WhatsApp* (FERRARI, 2019), podendo causar um impacto significativo de acordo com as estratégias implementadas no seu uso. Atualmente, o ser humano encontra uma necessidade de ser socialmente visíveis para os outros, esforço o qual na maior parte é descartado pelos outros que simplesmente recebem uma quantidade absurda de informações no feed de notícias do *Facebook*, por exemplo. Em vista disso, é observado o quão exposto um usuário de rede social pode colocar suas informações pessoais ao alcance de qualquer um.

Como as mídias têm esse poder de dispersão de informação em frações de segundo, juntamente com a necessidade de consumo dessas informações pelos seus milhões de usuários diários, surgem oportunidades de redução de esforço na publicidade de certas organizações, como por exemplo, as Instituições de Ensino Superior. Essa proximidade permite a conversão de usuários para alunos agregados à Instituição de maneira rápida e eficiente. Uma promoção em um curso específico, mostras de estrutura e capacitação dos profissionais fazem com que a pessoa sinta vontade de realizar um curso superior. Outro fator decisivo que faz com que a procura por uma faculdade são as exigências cada vez maiores dos mercados de trabalho, que procuram sempre os melhores e mais preparados profissionais para atuar em determinada área. Para isso, as Instituições investem pesado em áreas comerciais voltadas à divulgação em massa de seu produto, como por exemplo, o Marketing Digital, Publicidade e Propaganda Digitais e Engenharia de Redes Sociais.

Da mesma forma que a divulgação acontece, a instituição também está exposta ao negativismo gerado na internet, pois um usuário mal-intencionado ou insatisfeito pode expor sua experiência ruim para outros usuários e esses podendo ser persuadidos e desistir de sua escolha. Como não existe um controle sobre o que está na internet, podem existir diversos comentários, bons ou ruins, sobre determinado assunto ou acontecimento que pode provocar sentimentos na sociedade. Em vista disso, é importante analisar a quantidade e o teor desses comentários sobre tal assunto para talvez assim reproduzir novamente decisões corretas e evitar o prejudicial, a fim de gerar valor tanto para as instituições quanto para as pessoas que estão relacionadas à mesma, seja pela prestação de serviços como pelo consumo do mesmo.

Neste contexto e com o propósito de contribuir com a efetividade dessas técnicas de divulgação e dar suporte na análise de dados das instituições de ensino superior, propôs-se, neste artigo, o estudo de técnicas de análise de sentimento em comentários voltados a essas instituições em redes sociais visando identificar indicadores de satisfação e possíveis estratégias de divulgação para auxiliar nas tomadas de decisões e na aplicação de marketing direcionado.

Considerando o cenário apresentado, foram realizados os seguintes objetivos específicos:

- Obter informações de maneira legal: conseguir os dados de forma anônima que não infringe leis de direitos morais, autorais e quaisquer fatores que envolvam exposição indevida de opinião ou argumento.

- Implementar uma análise de sentimento: realizar a análise de sentimento com base nos dados obtidos anteriormente por meio de raspagem de dados e consultas em *API*'s públicas.
- Definir estratégias de exibição dos indicadores: quais informações serão relevantes para as instituições e como exibi-las de maneira intuitiva e eficiente.
- Criar um ambiente de análise: tratar os dados de maneira eficaz e gerar indicadores de qualidade para as instituições.

Portanto, de acordo com os pontos apresentados acima, a análise desses indicadores e dados gerados pela aplicação servirá de material de apoio na elaboração de planos estratégicos por parte das instituições de ensino para a obtenção de novos alunos e demais interessados na área.

2 REFERENCIAL TEÓRICO

Nesta seção são apresentados conceitos e estudos realizados sobre raspagem de dados e *webcrawling* para obtenção de dados, análise de sentimentos e marketing digital voltado às Instituições de Ensino Superior.

2.1 RASPAGEM DE DADOS E WEBCRAWLING

Data scraping (do inglês, raspagem de dados, comumente chamado de *web scraping*) é uma técnica computacional na qual um programa extrai dados de saída legível somente para humanos, proveniente de um serviço ou aplicativo. Teoricamente, *web scraping* é a prática de coletar dados por qualquer meio que não seja um programa interagindo com uma *API* (ou, obviamente, por um ser humano usando um navegador web). (MITCHELL, 2018). Normalmente, a transferência de dados é feita utilizando-se estrutura de dados adequada para processos automatizados por computadores, e não por humanos. Tais comunicações de formato e protocolo são rigidamente estruturadas, documentadas, facilmente analisadas, mantendo a ambiguidade ao mínimo. Primeiro, é coletada a página alvo em forma de texto não estruturado, segundo, é determinado seus padrões recorrentes por trás da informação a qual se deseja e por último são aplicados esses padrões no texto não estruturado para se obter as informações desejadas (MUNZERT, 2014).

A raspagem de dados é mais frequentemente realizada em serviços ou aplicativos web legados ou aqueles que não oferecem uma interface de programação de aplicação. Neste último caso, geralmente os responsáveis pelo serviço ou aplicação consideram a raspagem de dados como indesejada, possivelmente pela sobrecarga do sistema, perda de receita por propagandas, ou pela perda do controle do conteúdo da informação.

A extração de dados é muitas vezes considerada *ad hoc*, uma técnica deselegante, frequentemente utilizada como último recurso quando não há outro mecanismo de intercâmbio de dados. Páginas *Web* são construídas utilizando-se linguagens de marcação baseadas em textos (*HTML e XHTML*) e frequentemente contêm uma riqueza de dados textuais úteis. No entanto, a maioria das páginas da web são projetadas para usuários finais humanos e não para uso automatizado. Devido a

isso, criou-se ferramentas que raspam conteúdo da web. Um *web scraping* é uma *API* para extrair dados de um web site. Empresas como a *Amazon AWS* e *Google* fornecem ferramentas de extração, serviços e dados públicos disponíveis sem custo para os usuários finais. Novas formas de *web scraping* envolvem capturar feeds de dados de servidores *web*. Por exemplo, *JSON* é comumente usado como um mecanismo de armazenamento de transporte entre o cliente e o servidor *Web*.

Recentemente, as empresas têm desenvolvido sistemas web que dependem de técnicas em análise do Modelo de Objeto de Documentos, visão computacional e processamento de linguagem natural para simular o processamento humano que ocorre ao visualizar uma página da web para extrair automaticamente informações úteis. Outro ponto bastante analisado é a legalidade da extração dessas informações, que quando essa informação coletada é utilizada para fins pessoais, não existe problema aparente, mas, se essa informação for republicada ou chegar a cunhos públicos sem os devidos reconhecimentos autorais ou permissões, isso pode acarretar problemas legais dependendo do tipo desse conteúdo (LAWSON, 2015).

Portanto, a aplicação se beneficiou dessa técnica para obtenção dos comentários em redes sociais para a realização da análise de sentimento em cima desses comentários obtidos, de forma legal e sem a exposição dos usuários autores, utilizando bibliotecas desenvolvidas pela comunidade com a linguagem de programação *python*.

2.2 ANÁLISE DE SENTIMENTO

Análise de sentimentos tem se tornado um importante tópico na *Web*, especialmente em redes sociais, com o desenvolvimento de aplicações para monitoramento de produtos e marcas, assim como a análise da repercussão de eventos importantes (BENEVENUTO, RIBEIRO e ARAÚJO, 2015). Nesse contexto, surgiram sistemas para tratar opiniões automaticamente, utilizando-se dos conceitos da área de Análise de Sentimentos (AS), também conhecida como Mineração de Opiniões (SILVA, LIMA e BARROS, 2012).

A análise de sentimentos refere-se ao uso de processamento de linguagem natural, análise de texto, linguística computacional e biometria para identificar sistematicamente, extrair, quantificar e estudar estados afetivos e informações subjetivas com a tarefa de identificar se a opinião que foi expressa em um determinado texto é positiva ou negativa. A ascensão das mídias sociais, como blogs e redes sociais tem despertado interesse na análise de sentimento.

Com a proliferação de opiniões, avaliações, recomendações e outras formas de expressão on-line, a opinião se transformou em uma espécie de moeda virtual para empresas que desejam comercializar os seus produtos, identificar novas oportunidades e gerenciar suas reputações. Como as empresas estão cada vez mais interessadas em automatizar o processo de pesquisa sobre o que está sendo falado sobre elas, a compreensão das conversas nas mídias sociais, a fim de identificar conteúdo relevante, que permita tomar ações específicas, tem levado as empresas a buscar soluções de análise de sentimentos.

Portanto, a aplicação se beneficiou dessa tecnologia, disponibilizada pela *Azure Cognitive Services* para criar indicadores de satisfação e levantar dados importantes para as instituições, sendo aplicada para analisar os comentários obtidos.

2.3 MARKETING DIGITAL

“Marketing é a atividade humana dirigida para satisfazer necessidades e desejos por meio de troca”. Essa definição contém em si alguns dos aspectos essenciais do marketing, que é dirigido para satisfazer necessidades e desejos humanos e, dessa forma, precisa levar em consideração o público alvo antes de tudo, conhecendo-o para poder satisfazê-lo (GABRIEL, 2010, p. 28). O Marketing digital seria uma modalidade de marketing onde ações de comunicação são realizadas de modo que as empresas possam utilizar por meio da internet, da telefonia celular e outros meios digitais para divulgar e comercializar seus produtos, conquistando novos clientes e melhorando a sua rede de relacionamentos. Ele engloba a prática de promover produtos ou serviços pela utilização de canais de distribuição eletrônicos, para então chegar aos consumidores rapidamente de forma relevante, personalizada e eficientemente.

Esse tipo de marketing traduz-se em ações adaptadas aos meios digitais, de forma a obter, nestes canais, a mesma eficiência e eficácia do marketing direto e, simultaneamente, potencializar os efeitos do marketing tradicional. Na sua operacionalização são, normalmente, utilizados canais, meios e ferramentas digitais. Segundo Martha Carrer Cruz Gabriel (2008, Marketing de otimização de buscas na web), o termo “marketing digital” não existe. O que determinará o trajeto do marketing em si são as estratégias de uso de plataformas digitais ou tradicionais, sendo que a escolha por uma é definida no modo como as pessoas as utilizam.

Toda a construção de um plano de ação se inicia com uma visão, e, junto a ela, um objetivo a ser alcançado. A partir deste traço são definidos os meios a serem utilizados, para enfim atingir os resultados. O tratamento singular a cada cliente é o segredo para que a escolha de um plano de ação seja de fato apropriada, somente o conhecer de um cliente ou projeto em questão fará com que os esforços de mídia potencializem o trabalho de marketing, o tornando assim digital. Dessa forma, não é mais possível pensar em ações isoladas na internet ou nas redes sociais. Seja qual for seu negócio, com toda a certeza uma parcela significativa de seus consumidores é representada por usuários conectados que acessam a mídia. Assim, incluir estratégias digitais como parte de seu planejamento estratégico e de marketing, criando ações integradas que permitam utilizar melhor os recursos disponíveis (TORRES, 2018).

De acordo com o apresentado, a aplicação levanta dados e indicadores com os comentários coletados e analisados conforme dito nos tópicos 2.1 e 2.2 para o uso estratégico das instituições de ensino em pontos específicos de sua divulgação a fim de gerar novos LEADS.

2.4 CAPTAÇÃO DE ALUNOS PELAS INSTITUIÇÕES DE ENSINO (LEAD)

A captação de alunos consiste em uma gestão de estratégias de marketing, seja digital ou social, para conseguir alunos para determinados cursos e atividades realizadas pelas Instituições de Ensino. Essa prática tem como foco o uso de meios de comunicação para divulgação de eventos e promoções realizadas para a conversão em matrículas correntes. No início dos anos 90, surge a era do marketing de relacionamento, com esforço e ampliação do foco concentrado no cliente. A partir dessa era, é reconhecido o valor e o potencial de lucro na conservação dos clientes. Com o mercado inovando todos os dias, com os clientes exigindo cada vez mais preço e devido ao maior uso da tecnologia da informação na gestão dos negócios surgiu a necessidade de uma forte relação entre a empresa e o consumidor, para que os clientes continuem comprando (DA SILVA NILSEN e HÜSKES, 2008).

Em um primeiro momento, para as instituições que se viram obrigadas a estabelecer estratégias para evitar a diminuição do número de seus alunos, a noção de marketing envolvia basicamente a noção do senso comum: o marketing se caracterizava como venda ou propaganda (COLOMBO, 2005). Mas, com o passar do tempo, essas mesmas instituições viram que seria possível realizar ações para que, ao invés de tentar manter seus clientes, identificar potenciais clientes para novos contratos, estratégia a qual é denominada "lead". O lead é um potencial cliente que, de alguma forma, demonstrou interesse no produto ou serviço que a instituição oferece, seja entrando em contato para obter mais informações, acessando perfis nas redes sociais ou nos portais de ensino, entre outros, despertando o interesse por diversos motivos: formação profissional, recolocação no mercado, aquisição de novos conhecimentos e excelência acadêmica. Essa prática permite que mais alunos interessados se matriculem nas instituições e assim manter uma taxa alta de clientes, mesmo perdendo alguns mais antigos.

Com isso, temos uma relação entre o Marketing Digital, citado anteriormente, que visa dar suporte à geração de leads da instituição por meio da publicidade e estratégias de obtenção, onde a aplicação tem um importante papel em apontar em quais pontos a comunidade tem mais atenção nas instituições de ensino e que se focados, conseguem um nível maior de captação com menor recurso.

3 METODOLOGIA

A aplicação foi desenvolvida tendo por base o estudo de técnicas de análise de sentimento e coleta de dados na *web* por meio de raspagem de dados, além de pesquisas relacionadas a estratégias de divulgação, mídias e redes sociais. Após a realização das pesquisas, primeiramente foi desenvolvida uma *API* para coleta de dados (comentários nas redes sociais) na linguagem *Python*, linguagem de programação de alto nível, interpretada, de script, imperativa, orientada a objetos, funcional. Os dados coletados nessa *API* foram armazenados no MongoDB, um banco de dados de alta performance orientado a documentos e não relacional.

Em sequência foi desenvolvida uma *API* para a gestão desses dados utilizando *NodeJS*, interpretador de JavaScript assíncrono com código aberto orientado a eventos,

focado em migrar a programação do *JavaScript* do cliente (*frontend*) para os servidores, criando aplicações de alta escalabilidade (como um servidor *web*), manipulando milhares de conexões/eventos simultâneas em tempo real numa única máquina física. Essa *API* é responsável por realizar chamadas à *API Text Analysis* provida pela *Azure Cognitive Services*, gerando indicadores e dados relacionados ao processo de análise de sentimentos dos comentários coletados. Após isso, foi desenvolvida a interface da aplicação em *VueJS*, biblioteca *JavaScript* de código aberto com foco em criar interfaces de usuário (*frontend*) em páginas *web*. Na conclusão da conexão entre os módulos desenvolvidos, a aplicação é hospedada no *Heroku*, plataforma em nuvem como um serviço que suporta várias linguagens de programação.

4 DESENVOLVIMENTO E RESULTADOS

A princípio foram realizadas pesquisas sobre as redes sociais, o impacto que elas trazem aos usuários e qual a frequência de utilização das pessoas sobre esses meios de comunicação e interação social, para determinar quais ambientes seriam o foco da coleta de comentários e materiais para a análise da aplicação. Também foram realizados estudos nas tecnologias selecionadas para a criação dos componentes visuais, dos módulos de análise e dos serviços que são consumidos e utilizados para a gestão dos dados, desde o processamento até o envio para exibição, além da análise da estrutura das páginas das redes sociais na internet para a realização da coleta dos dados.

Após as pesquisas e estudos, foram criados os módulos de coleta de comentários nas redes sociais, utilizando *Python* e técnicas de *webscraping* e *webcrawling*, que permitem a automatização da procura e da obtenção de resultados, análise de conteúdos de páginas web e assim a coleta de dados que são relevantes para a aplicação como um todo. Na Figura 1, é apresentado um trecho de código que é responsável pela coleta automática de comentários em uma rede social.

Figura 1: Código de coleta de comentários no Twitter

```
modules | helloworld | @ helloworld
12 def twitterSearch(query, maximum=10):
13     options = Options()
14     options.headless = True
15     options.add_argument('--window-size=1920,1200')
16     options.add_argument('--disable-dev-shm-usage')
17     options.add_argument('--no-sandbox')
18
19     if os.environ.get('ENV') == 'development':
20         driver = webdriver.Chrome()
21         ChromedriverManager().install(), options=options)
22     else:
23         options.binary_location = os.environ.get('GOOGLE_CHROME_BIN')
24         driver = webdriver.Chrome(executable_path=os.environ.get(
25             'CHROMEDRIVER_PATH'), chrome_options=options)
26
27     url = 'https://twitter.com/search?q=' + query + '&src=recent_search_click&live'
28
29     finalComments = []
30     # finalComments = []
31
32     try:
33         driver.get(url)
34         time.sleep(2)
35
36         while True:
37             javascript = "window.scrollTo(0, document.body.scrollHeight);"
38             driver.execute_script(javascript)
39             time.sleep(2)
40
41             element = driver.find_element_by_tag_name("body")
42             commentsDiv = element.get_attribute("innerHTML")
43
44             soup = BeautifulSoup(commentsDiv, 'html.parser')
45
46             comments = soup.find_all('div', attrs={
47                 'class': 'css-901oao r-1k9yb0a r-1q8b0m r-ab22e0 r-16dbaf1 r-ad920e r-bc0qeo r-bwqia r-quot0'})
48
49             for i in comments:
50                 if i in finalComments:
51                     continue
52                 else:
53                     finalComments.append(html.escape(i.text))
54
55             if len(finalComments) >= maximum:
56                 break
57
58     print('twitter crawling done!')
```

Fonte: Dados da pesquisa, 2020.

Após finalizar a coleta dos dados, os mesmos são armazenados em um banco de dados não relacional *MongoDB*, que é capaz de armazenar estruturas de dados de maneira documental e que apresenta alta performance no acesso e na manipulação de seus documentos. Com os dados armazenados, foi criada uma *API* em *NodeJS*, seguindo o padrão *REST* para comunicação com o banco de dados e para o envio dos comentários para a análise de sentimento, de acordo com os padrões de uso da Microsoft e o Azure Cognitive Services. Na Figura 2, é apresentado o trecho de código responsável pela obtenção dos documentos no banco de dados, na organização desses nos padrões e a requisição à Azure para receber a análise de sentimento.

Figura 2: Código de requisição à Azure para análise dos dados

```
src > controllers > analysisData.js > (6) index
You, a few seconds ago | 1 author (You)
1  "use strict";
2
3  const Data = require("../schemas/dataSchema");
4  const azureUtil = require("../utils/azure.util.js");
5  const returnUtil = require("../utils/return.util.js");
6
7  const index = async (req, res) => {
8    try {
9      const id = req.query.id;
10
11     const result = await Data.findOne({ _id: id }, async function (err, result) {
12       if(err) { return; }
13
14       return result;
15     });
16
17     const documents = azureUtil.buildRequest(result.comments);
18     const analysedData = await azureUtil.makeRequest(documents);
19     const analysis = returnUtil.buildAnalysisObject(
20       result,
21       documents,
22       analysedData
23     );
24
25     result.analysis = analysis;
26
27     console.log(result.analysis);
28
29     await Data.updateOne({ _id: id }, { analysis: analysis });
30
31     res.status(200).json({ message: "OK!" });
32   } catch(error) {
33     res.status(500).json({ error: "Error: " + error });
34   }
35 };
36
37 module.exports = {
38   index,
39 };
40
```

Fonte: Dados da pesquisa, 2020.

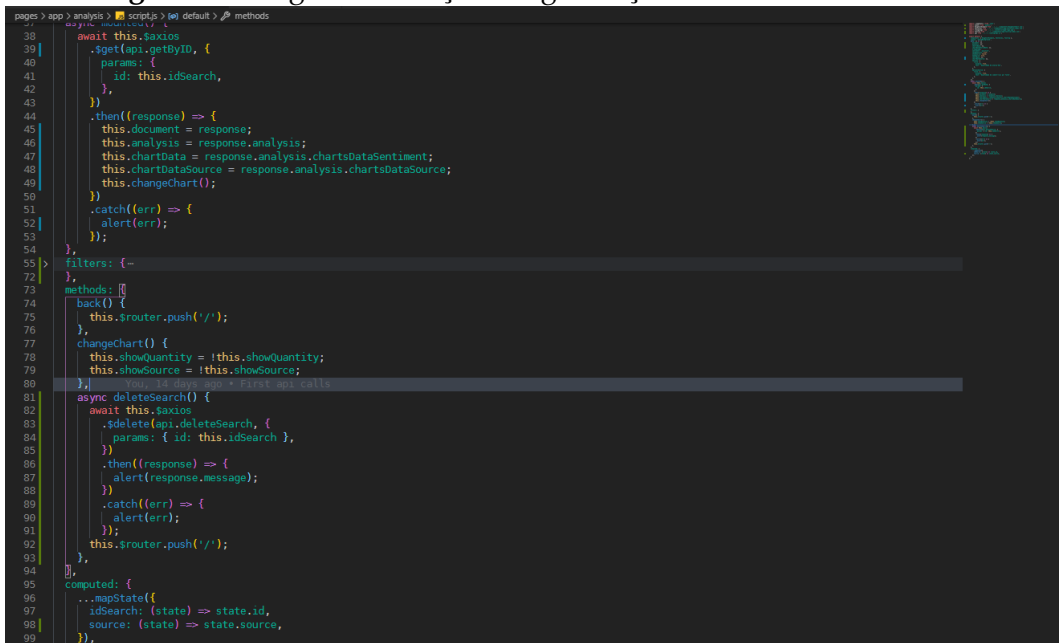
Assim que a análise é finalizada, os métodos de análise são responsáveis por gerar as estatísticas que serão consumidas pela interface da aplicação para a exibição dos indicadores, sendo realizados cálculos matemáticos simples, como médias, porcentagens e contagem dos dados, além de organizar dados relevantes, como sentenças que impactaram na análise de cada comentário e seus *scores* de avaliação.

Feito isso, o documento contendo os comentários é manipulado e assim é enviado para que a interface da aplicação consiga exibir os resultados da análise. Na interface de exibição, os dados são coletados por meio de requisições também no padrão *REST*, sendo manipulados pelos *scripts* em *JavaScript* interpretados pelo *VueJS*, organizados em componentes e propriedades que são persistidos no estado da aplicação para exibição e interação com o usuário. Na Figura 3 é apresentado um trecho do código de obtenção e organização desses dados.

Figura 3: Código de obtenção e organização dos dados da interface

```
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```



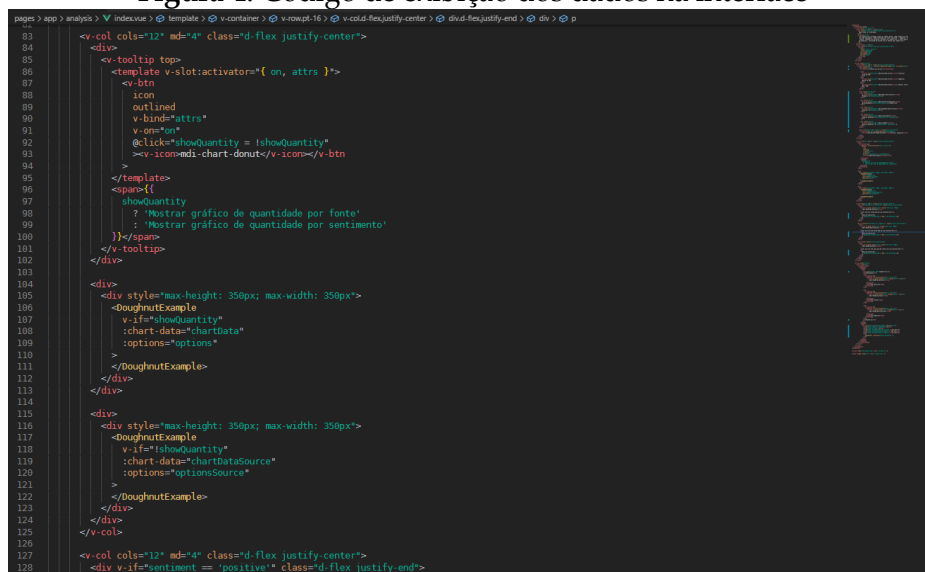
Fonte: Dados da pesquisa, 2020.

Com os dados organizados, o *VueJS* também é capaz de interpretar linguagem de marcação (*HTML*, *CSS*), com algumas propriedades nativas do *framework* de desenvolvimento *front-end*, permitindo a exibição dos dados de maneira limpa e organizada, juntamente com a possibilidade de configuração de visualização responsiva (dispositivos *mobile*) de maneira ágil e simplificada. Na Figura 4, uma representação dessa organização de componentes que serão visualizados em código *VueJS*.

Figura 4: Código de exibição dos dados na interface

```
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128

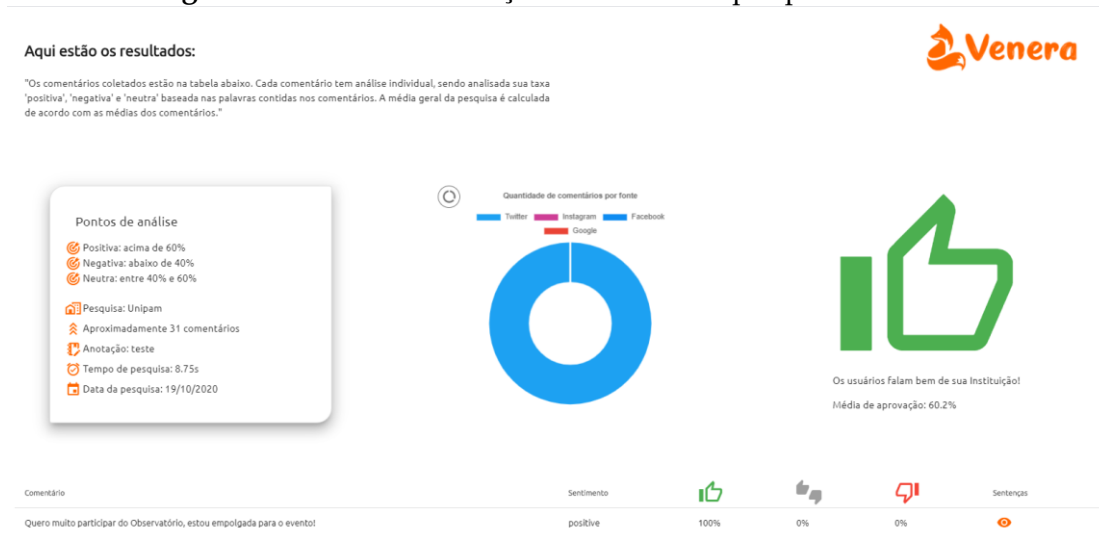
```



Fonte: Dados da pesquisa, 2020.

O resultado é a interface totalmente populada com os dados disponibilizados pela análise, processados pelos módulos de comunicação e conexão juntamente com a Azure, sendo apresentados de maneira intuitiva e simples, como apresentado na Figura 5.

Figura 5: Interface de exibição dos dados de pesquisa e análise



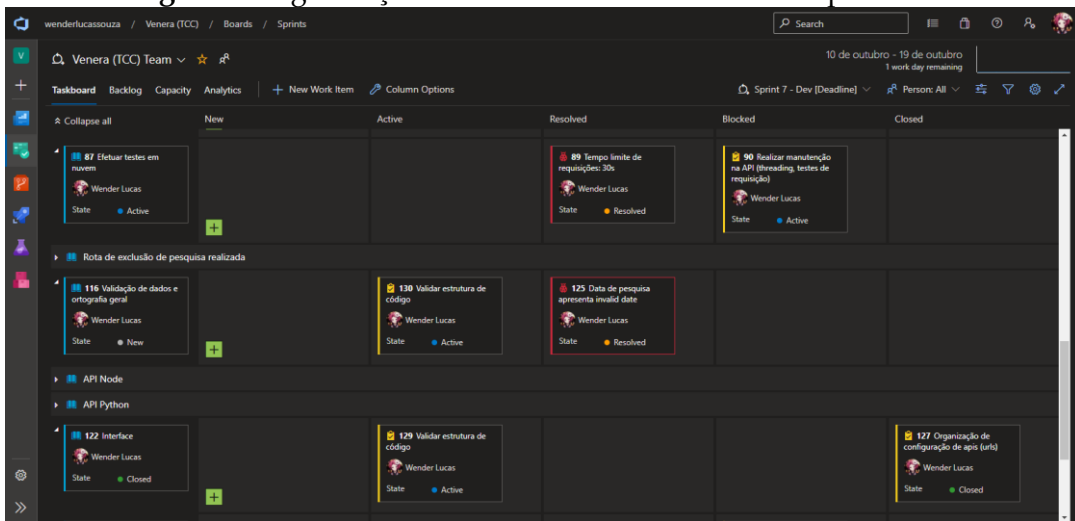
Fonte: Dados da pesquisa, 2020.

Na Figura 5, são apresentados os seguintes elementos:

- Pontos de análise e dados de pesquisa: São apresentados os pontos principais de análise da aplicação e os dados que foram realizados na pesquisa;
- Gráficos quantitativos: São apresentados os gráficos quantitativos por fonte de obtenção do comentário (rede social utilizada) e também por análise (positiva, negativa e neutra);
- Média geral de aprovação: É apresentada a média geral de aprovação de acordo com a análises dos comentários;
- Tabela de comentários: São apresentados todos os comentários utilizados na análise, seus *scores* individuais, as sentenças que mais impactaram na análise e o sentimento do comentário.

As tarefas realizadas durante o desenvolvimento da aplicação foram devidamente registradas e organizadas de acordo com o padrão de desenvolvimento ágil *SCRUM*, sendo utilizado o Azure DevOps para gerência das atividades, como mostrado na Figura 6.

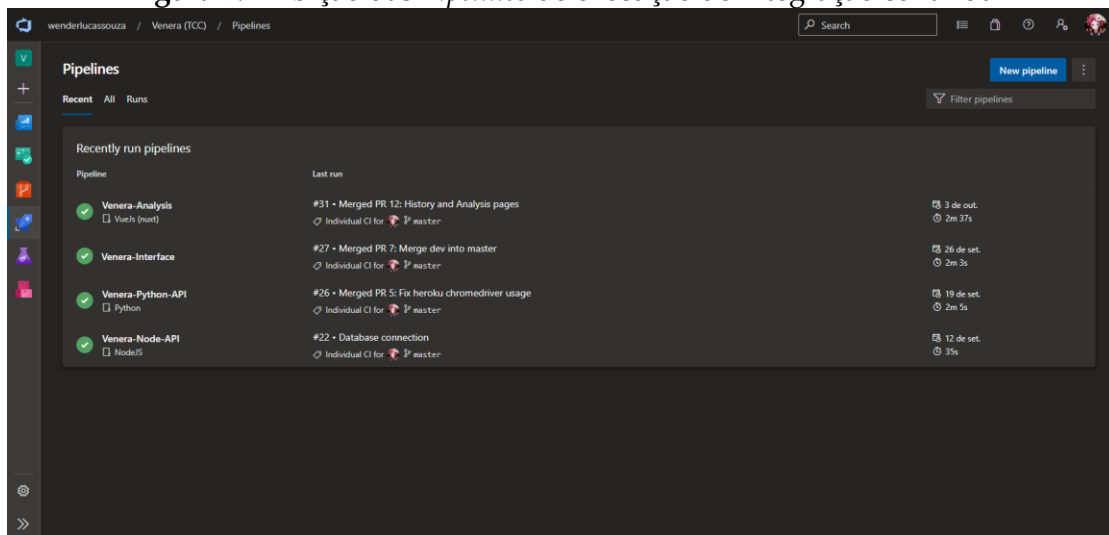
Figura 6: Organização das atividades no Azure DevOps Boards



Fonte: Dados da pesquisa, 2020;

Por fim, para que a aplicação fosse disponibilizada para uso, foram criadas *Pipelines* de execução de integração contínua com os serviços em nuvem para *deploy* e hospedagem dos repositórios contendo as aplicações. Para isso, foram configuradas essas *Pipelines* no Azure DevOps, integradas com o serviço de *cloud Heroku* e assim, basta acessar a aplicação pelo domínio disponibilizado na conclusão da integração. Nas Figuras 7 e 8 são exibidas as *Pipelines* e as execuções.

Figura 7: Exibição das *Pipelines* de execução de integração contínua

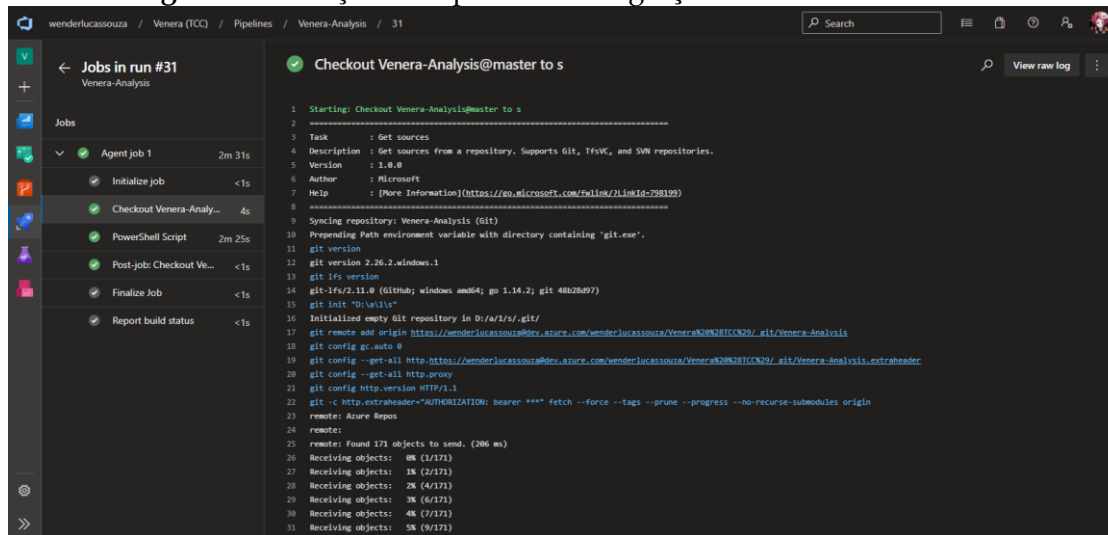


Fonte: Dados da pesquisa, 2020.

Como exibido na Figura 7, cada repositório contém suas *Pipelines* de execução independente, fazendo com que cada módulo do sistema consiga ser disponibilizado separadamente, porém se integrando internamente com os outros módulos para o funcionamento da aplicação como um todo. Na Figura 8, é apresentado um exemplo de execução de uma *Pipeline*, contendo informações de disponibilização, versão do

repositório e os comandos necessários para que a integração seja feita de maneira correta.

Figura 8: Execução da Pipeline de integração contínua da interface



Fonte: Dados da pesquisa, 2020.

De acordo com o sucesso das integrações e da implantação da aplicação em nuvem, a plataforma é disponibilizada para uso aberto.

5 CONCLUSÃO

O estudo demonstrado objetivou coletar comentários de usuários nas redes sociais, aplicar técnicas de Inteligência Artificial por meio de Análise de sentimentos em textos e converter os resultados das análises em indicadores de satisfação, dados estatísticos e expor a opinião desses usuários sobre determinados assuntos, principalmente voltado à instituições de ensino superior para que essas possam agir em cima desse feedback para a captação de novos alunos ou na melhoria de estratégias de marketing empresarial, social e geral.

Como dito anteriormente, o sistema desenvolvido foi projetado voltado para as instituições de ensino superior, porém essa proposta poderá se expandir para diferentes áreas, tanto no setor de educação quanto em outros setores gerais, os quais possam se beneficiar da análise de comentários nas redes sociais sobre determinada marca, assunto ou discussão.

Os objetivos deste artigo foram considerados alcançados parcialmente, pois com a execução de todas as etapas, foi conseguido o levantamento de dados estatísticos que comprovem o nível de satisfação/insatisfação relacionado a instituições de ensino, juntamente com opiniões sobre as mesmas, porém, com a visão de que o sistema possa atender a outras áreas, uma gama de possibilidades é criada a partir do conceito de que todo e qualquer comentário em uma rede social pode ter seu sentimento analisado e que esse ato pode gerar um indicador de satisfação relacionado a pesquisa realizada.

Tendo em mente o atual de mercado de trabalho, as exigências que o mesmo pede para que sempre sejam conseguidos os melhores profissionais nas mais diversas áreas existentes, o Ensino Superior e seu contexto social estão sempre em pauta, ao passo que as pessoas também procuram as melhores instituições para aprendizado e preparação acadêmica, e utilizando as ferramentas que a tecnologia nos propões nos dias de hoje, como por exemplo a ascendente Inteligência Artificial, para nos auxiliar no processo de unificação entre pessoas que buscam o conhecimento e a busca por profissionais de qualidade, o sistema poderá ser de grande utilidade nessa empreitada tanto para as Instituições de Ensino Superior quanto para seus possíveis clientes.

Para melhor aproveitamento e utilização do sistema em momentos futuros, a exploração de novas áreas além da educação superior é um diferencial que pode tornar a ferramenta ainda mais poderosa, a fim de gerar dados estatísticos a partir de qualquer problema ou assunto que se deseja realizar a pesquisa. Também, com o avanço dos estudos no campo da inteligência artificial ou do desenvolvimento voltado à web, a plataforma pode se tornar ainda mais eficiente e performática de acordo com a aplicação de outras técnicas de coleta e análise de sentimento, aumentando assim a precisão e a veracidade da análise, gerando dados ainda mais reais sobre a pesquisa e aumentando também a qualidade da tomada de decisões baseada no levantamento desses indicadores.

REFERÊNCIAS

AGUIAR, Sonia. **Redes sociais na internet: desafios à pesquisa**. In: XXX Congresso Brasileiro de Ciências da Comunicação. 2007.

BENEVENUTO, Fabrício; RIBEIRO, Filipe; ARAÚJO, Matheus. **Métodos para análise de sentimentos em mídias sociais**. In: Brazilian Symposium on Multimedia and the Web. (Webmedia), Manaus, Brazil. 2015. p. 30.

CAMACHO, Dóris. **Marketing digital: compra de mídia e inbound**. São Paulo: Editora Senac São Paulo, 2019.

COLOMBO, Sonia Simões. **Marketing Educacional em Ação: estratégias e ferramentas**. Porto Alegre: Bookman Editora, 2005.

DA SILVA NILSEN, Andréia; HÜSKES, Monika Ilse. Marketing de relacionamento para a fidelização e captação de alunos na Wizard Blumenau. **Revista Interdisciplinar Científica Aplicada**, v. 1, n. 2, p. 1-21, 2008.

FERRARI, Pollyana. **A força da mídia social: interface e linguagem jornalística no ambiente digital**. [S. l.]: Editora Estação das Letras e Cores, 2019.

GABRIEL, Martha Carrer Cruz. **Marketing de otimização de buscas na web**. São Paulo: Ed. Esfera, São Paulo, 2008.

GABRIEL, Martha Carrer Cruz. **Marketing na era digital**: conceitos, plataformas e estratégias. São Paulo: Novatec Editora, 2010.

LAWSON, Richard. **Web scraping with Python**. Packt Publishing Ltd, 2015.

MITCHELL, Ryan. **Web scraping with Python**: collecting more data from the modern web. "O'Reilly Media, Inc.", 2018.

MUNZERT, Simon et al. **Automated data collection with R: A practical guide to web scraping and text mining**. John Wiley & Sons, 2014.

SILVA, Nelson Rocha; LIMA, Diego; BARROS, Flávia. **Sapair**: um processo de análise de sentimento no nível de característica. *In: 4nd International Workshop on Web and Text Intelligence (WTI'12)*, Curitiba. 2012. p. 2.

TORRES, Cláudio. **A bíblia do marketing digital**: tudo o que você queria saber sobre marketing e publicidade na internet e não tinha a quem perguntar. São Paulo: Novatec Editora, 2018.

ZIMMERMAN, Jan. **Marketing digital para leigos**. Tradução de Marcella de Melo Silva São Paulo: Alta Books Editora, 2018.