# SIMULADOR DE POSTAGENS NO INSTAGRAM PARA AUMENTO DO ENGAJAMENTO: APLICAÇÃO NO SETOR DE FOOD SERVICE

Kellen Dayelle Endler - Universidade Federal do Paraná - UFPR

Claudio Luis Cruz De Oliveira - Escola Superior de Propaganda e Marketing - ESPM

Carolina Toledo Ferraz - Escola Superior de Propaganda e Marketing - ESPM

Letícia Duarte Jimenez - Escola Superior de Propaganda e Marketing - ESPM

**Resumo**

O engajamento nas redes sociais é essencial para empresas de Food Service, pois contribui para o estabelecimento de relacionamentos com os clientes e amplia a visibilidade da marca. O Instagram, por sua vez, dentre as redes sociais, se tornou um canal essencial para promover estabelecimentos de comida, devido à sua abordagem visual, alcance amplo e interação direta com os clientes. Este artigo apresenta o estudo sobre o desenvolvimento de um simulador de postagens no Instagram aplicado a esse setor. O objetivo principal desse simulador consiste em otimizar estratégias de postagens para aumentar o número de curtidas em publicações de estabelecimentos de alimentação. O simulador incorpora variáveis como características da presença digital da empresa, seu número de publicações no perfil, número de seguidores, formatos da publicação, legendas e hashtags. Este estudo visa preencher a lacuna da falta de ferramentas específicas para esse setor, inovando ao fornecer uma ferramenta prática para maximizar curtidas. Essa contribuição não apenas auxilia as empresas a maximizar sua visibilidade e atratividade online, como também promove a eficácia no uso de recursos de marketing. Por fim, os resultados obtidos podem contribuir com pesquisas futuras e aprimoramento contínuo de táticas de engajamento em redes sociais.

**Palavras-chave:**Serviços de Alimentação; Inteligência Artificial; Instagram.

**Abstract**

Engagement on social media is essential for Food Service companies, as it contributes to establishing relationships with customers and increases brand visibility. Instagram, among social networks, has become an essential channel for promoting food establishments, due to its visual approach, broad reach and direct interaction with customers. This article presents the study on the development of an Instagram posting simulator applied to this sector. The main objective of this simulator is to optimize posting strategies to increase the number of likes on publications from food establishments. The simulator incorporates variables such as characteristics of the company's digital presence, its number of profile publications, number of followers, publication formats, captions and hashtags. This study aims to fill the gap in the lack of specific tools for this sector, innovating by providing a practical tool to maximize likes. This contribution not only helps companies maximize their online visibility and attractiveness, but also promotes effective use of marketing resources. Finally, the results obtained can contribute to future research and continuous improvement of engagement tactics on social networks.

**Keywords:** Food Services; Artificial intelligence; Instagram.

# SIMULADOR DE POSTAGENS NO INSTAGRAM PARA AUMENTO DO ENGAJAMENTO: APLICAÇÃO NO SETOR DE *FOOD SERVICE*

## 1 INTRODUÇÃO

Segundo a Associação Brasileira de Bares e Restaurantes (ABRASEL), o segmento de franquias de alimentação teve um faturamento de R$ 51 bilhões em 2022, um aumento de 18% em relação a 2021, com crescimento de 9% no número de unidades. O setor cresceu 20% no primeiro trimestre de 2023 (ABRASEL, 2023a). Antes da pandemia, em 2018, o setor de alimentação fora do lar apresentou um faturamento de 230 bilhões de reais e já representa 34% dos gastos com alimentação dos brasileiros (ABRASEL, 2019). Já o Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (SEBRAE) revelou que o setor gerou 6 milhões de empregos em 1 milhão de negócios e representou 2,7% do PIB brasileiro em 2018 (SEBRAE, 2019). Esses números evidenciam a importância do setor e apontam para perspectivas promissoras de crescimento futuro.

Apesar da importância deste setor, existem a falta de ferramentas específicas e práticas que auxiliem essas empresas com desafios relacionados à sua presença digital. Considerando um panorama nacional, em 2020, houve um investimento de R$ 49 bilhões em campanhas publicitárias (KANTAR IBOPE MEDIA, 2021). Porém, o desperdício de verba publicitária pode chegar a mais da metade do total investido (CHAIT, 2015; FOU, 2020; THINKBOX, 2018). Esse é um problema antigo da publicidade, John Wanamaker, fundador da primeira loja de departamentos americana, afirmava que “metade do investimento em publicidade é desperdício; o problema é que não sei qual é essa metade” (CHAIT, 2015). Pesquisas indicam que a publicidade melhora a competitividade da economia, pois incentiva a diferenciação pela inovação, promove a competição pela qualidade e preço e estimula o crescimento do mercado (ABAP, 2015). Por isso, a otimização da verba publicitária é fundamental para o crescimento de nossa economia.

Segundo pesquisa realizada pela Associação Brasileira de Bares e Restaurantes (ABRASEL) com mais de 1.500 empresários de todo o Brasil, quase a totalidade dos estabelecimentos (99%) possui presença em, pelo menos, uma rede social para fins de *marketing* e relacionamento com os clientes (ABRASEL, 2023b). Sendo que as plataformas mais utilizadas são Instagram, WhatsApp e Facebook, nesta ordem. Esse fato evidencia que esse setor tem encontrado nas plataformas de mídia social, como o Instagram, um canal estratégico para atrair e engajar clientes.

O estudo global *State of Inbound Marketing Trends* (2022) mostra que as companhias que priorizam o *marketing* digital geram 2,8 vezes mais crescimento de receita em relação àquelas que não o fazem. Sendo importante destacar que o sucesso do crescimento de uma empresa de *Food Service* que utiliza as redes sociais dependerá da qualidade do conteúdo, engajamento com a comunidade e consistência nas postagens. Assim, existe a necessidade imperativa por parte das empresas de alimentação de se adotar estratégias inovadoras e aproveitar as últimas tecnologias para se destacarem.

O uso da inovação e tecnologia, aliado à inteligência artificial, representa um avanço no campo da análise de dados e estratégias de mídia social. A capacidade de gerar um modelo preditivo de engajamento com base no número de *likes* em postagens é um exemplo desse progresso. Por meio da coleta e análise de grandes volumes de dados de interação dos usuários em mídias sociais, a inteligência artificial é capaz de identificar padrões complexos e correlações sutis que escapariam à percepção humana. Essas descobertas alimentam a construção de modelos que podem prever com o impacto de diferentes variáveis, como conteúdo da postagem, o número de publicações no perfil da empresa, o número de seguidores,

os formatos das publicações, as legendas e as *hashtags* utilizadas, no nível de engajamento. Esse conhecimento não apenas auxilia as empresas a otimizar suas estratégias de postagens, mas também representa uma evolução na maneira como a tecnologia é aplicada para alcançar melhores resultados nas redes sociais.

Este artigo apresenta um estudo que visa abordar uma questão importante para as empresas de alimentação que buscam destacar-se nas redes sociais: como otimizar estratégias de postagens para aumentar o número de curtidas em suas publicações no Instagram. O foco principal deste trabalho é apresentar o desenvolvimento de um simulador de postagens no Instagram, que consiste em uma ferramenta inovadora e prática projetada para auxiliar empresas de *Food Service* a melhorar seu desempenho nas redes sociais, aumentando suas vendas.

O presente artigo está estruturado da seguinte forma: após a contextualização do tema, a seção 2 apresenta o problema de pesquisa e o objetivo do artigo. A seção 3 apresenta a fundamentação teórica, seguida da exposição detalhada da metodologia utilizada apresentada na seção 4. Os resultados obtidos são discutidos na seção 5, e, finalmente, na seção 6, é apresentada a conclusão do estudo, destacando suas implicações e possíveis direções para futuras pesquisas.

## 2 PROBLEMA DE PESQUISA E OBJETIVO

O problema central que este estudo se propõe a abordar é o desafio enfrentado pelas empresas de alimentação em otimizar suas estratégias de postagens para aumentar o número de curtidas em suas publicações. Apesar da importância das redes sociais para a visibilidade e o sucesso desses negócios, muitas empresas carecem de orientações precisas sobre como melhorar seu engajamento e atrair a atenção de seu público-alvo.

O objetivo principal desta pesquisa é apresentar o desenvolvimento de um simulador de postagens para o Instagram direcionado ao setor de *Food Service*, de modo a otimizar o engajamento avaliado pelo número de curtidas. Esse simulador integra inovação e tecnologia para solucionar o problema mencionado. Para atingir esse objetivo, os seguintes objetivos específicos foram definidos:

- Desenvolver o Simulador: Criar uma ferramenta tecnológica inovadora, por meio da inteligência artificial que seja capaz de analisar e simular estratégias de postagens no Instagram para empresas de alimentação.
- Disponibilizar a ferramenta para testes na internet: Criar uma interface de usuário acessível e amigável para que as empresas possam testar o simulador de forma prática e integrá-lo às suas operações de *marketing* digital.

Ao alcançar esses objetivos, este estudo busca contribuir significativamente para o setor de *Food Service*, fornecendo às empresas uma ferramenta inovadora e tecnologicamente avançada que as auxilie na maximização do impacto de suas estratégias de mídia social. Espera-se que o simulador desenvolvido seja uma solução prática para o problema de otimização de postagens no Instagram, permitindo que as empresas alcancem um maior engajamento do público e, por consequência, um aumento em sua visibilidade e sucesso no mercado digital.

## 3 FUNDAMENTAÇÃO TEÓRICA

A competição intensa e o compartilhamento constante de opiniões e fotos sobre comida tornaram o setor de restaurantes desafiador, exigindo que os gestores repensem suas estratégias e uso das mídias sociais (TUSSYADIAH E FESENMAIER, 2009). Apesar da relevância do uso das mídias sociais, as organizações ainda não lucram plenamente com eles (DOSSENA, et al., 2021). Isto deve-se principalmente à falta de recursos (por exemplo, tempo) e de

capacidades e competências necessárias para utilizar eficazmente as diferentes plataformas de acordo com os objetivos organizacionais acompanhando a sua constante transformação (ROY E DIONNE, 2016).

As competências de modelagem/análise e o pensamento crítico são também fundamentais para a integração do conhecimento compartilhado em redes sociais aos processos de negócio, pois permitem às organizações compreenderem e interpretar tendências e utilizá-las aos negócios (MCCLENEGHAN, 2006; TENCH E MORENO, 2015; D'ANGELO, 2010). Essas habilidades analíticas requerem criatividade, curiosidade intelectual e flexibilidade para interpretar tendências de forma original, juntamente com um conhecimento do negócio e rigor metodológico para aproveitar novos conhecimentos de maneira eficaz. A utilização estratégica das mídias sociais pode antecipar mudanças e permitir ajustes nos modelos de negócios e reconfiguração e integração de capacidades atuais para responder às oportunidades e ameaças que surgem no ambiente das mídias sociais (DOSSENA, et al., 2021).

Uma das formas de aproveitar melhor o poder das mídias sociais consiste em utilizar a Inteligência Artificial (IA) (SADIKU, 2021). Atualmente, as mídias sociais estão sendo usadas para inferir o comportamento social e derivar tendências, em combinação com ferramentas de análise de *big data* (SARMIENTO, 2020). A inteligência artificial, tem promovido, em muitos casos, resultados mais rápidos e mais assertivos comparados a quaisquer previsões precedentes (KAUFMAN, SANTAELLA, 2020). Bag et al. (2022) afirmam que a implantação de tecnologias de IA tem uma relação positiva com o envolvimento e a conversão do usuário. Além disso, a conversão leva a uma experiência do usuário satisfatória.

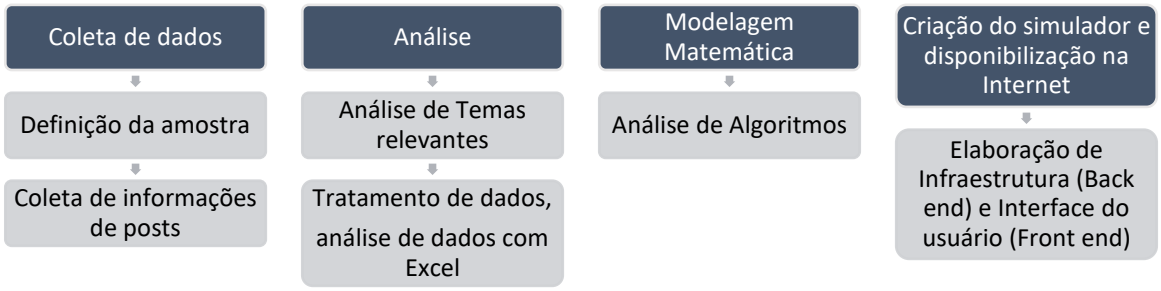
O Aprendizado de Máquina é uma técnica fundamental da Inteligência Artificial, que inclui uma ampla gama de algoritmos e modelos estatísticos que permitem que os sistemas encontrem padrões, façam inferências e aprendam a executar tarefas sem instruções específicas de forma autônoma (GREENGARD, 2019). Muitos são as aplicações dessas técnicas relacionadas a mídias sociais. Como por exemplo: para analisar grandes volumes de informações de conteúdos de redes sociais durante crises de saúde humanitária, visando identificar dados relevantes (FERNANDEZ-LUQUE, IMRAN, 2018); para analisar dados de mídias sociais para triagem de depressão (OWUSU, et al., 2021); para compreender as percepções públicas dos aplicativos de rastreamento de contatos COVID-19 (CRESSWELL, 2021), dentre outros.

Em síntese, a revisão da literatura apresentada neste artigo demonstra a crescente importância das redes sociais. Sendo que a aplicação proposta neste artigo, consiste na integração da inteligência artificial e da análise de dados como um elemento-chave para a otimização das estratégias de postagens, promovendo uma abordagem orientada por dados. Por meio da inteligência artificial, busca-se fornecer diretrizes práticas para empresas de alimentação que buscam destacar-se nas redes sociais, impulsionando o número de curtidas em suas publicações.

## 4 METODOLOGIA

A metodologia empregada envolveu etapas de coleta de dados, análise, elaboração de modelagem matemática e criação do simulador com sua disponibilização na internet. A Figura 1 apresenta a estrutura metodológica adotada juntamente com as etapas específicas realizadas em cada fase do processo.

Figura 1. Metodologia adotada




Fonte: Os autores (2023)

Na sequência, apresenta-se de forma detalhada as etapas da pesquisa. Cada etapa é crucial para a validade e confiabilidade dos resultados.

## 4.1 COLETA DE DADOS

A base de dados foi construída usando a API do Instagram chamada *Instaloader* (Disponível em: https://instaloader.github.io/). Foi desenvolvido um código em Python com o objetivo de usar a API e baixar as métricas necessárias das 28 empresas selecionadas como objeto de estudo, para o desenvolvimento do projeto. As empresas foram selecionadas com base no ranking das "300 Maiores Empresas do Varejo Brasileiro 2022" da (SBVC, 2022). Além dessas, outras franquias importantes do interesse dos pesquisadores foram adicionadas à amostra, considerando-se critérios como a diversidade de segmentos dentro do setor de *Food Service*, o tamanho das empresas e a presença digital delas. Os perfis do Instagram das 28 empresas analisadas são os seguintes:

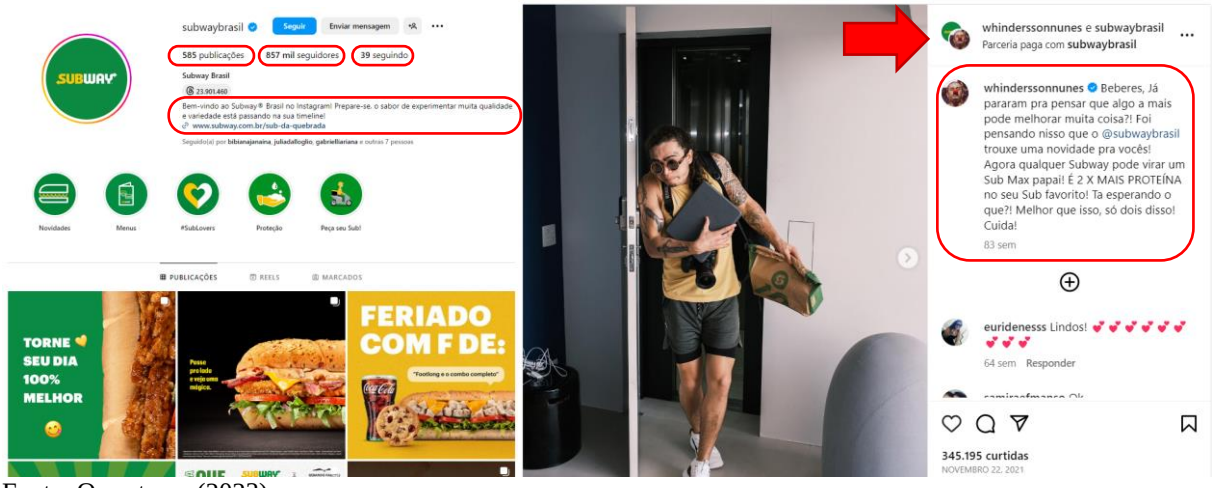
- bobsbrasil
- burgerkingbr
- cacaushow
- cafecultura
- chinainboxoficial
- chiquinhosorvetes
- croasonho
- didiodelivery
- dominospizzabrasil
- espettocarioca
- grilettooficial
- habibsoficial
- jinjinoficial
- kfcbrasil
- konistore
- lebonton.cuisine
- maderobrasil
- mcdonalds_br
- montanagrilloficial
- pizzahutbrasil
- redegiraffas
- reidomateoficial
- rest_bonaparte

- salebrasaoficial
- sodiedoces
- spoleto_oficial
- subwaybrasil
- sushichicoficial

Após a seleção das empresas, procedeu-se a coleta de dados propriamente dita, que envolveu a obtenção de informações dos perfis dessas empresas no Instagram. Ao todo, foram coletadas informações de 43.975 posts do Instagram provenientes do perfil dessas empresas. Esses posts coletados referem-se a todas as publicações dessas empresas desde as suas primeiras publicações até maio de 2023. As informações coletadas do Instagram incluíram informações como: número de publicações que a empresa publicou, contagem de seguidores, contagem de perfis que a empresa segue, descrição da bio, número de comentários, número de curtidas, descrição da legenda, vídeo, data de publicação, *hashtag* da legenda, menções da legenda e *url* da postagem para que os pesquisadores tivessem acesso fácil a consulta das publicações. É importante mencionar que nem todas as informações coletadas foram consideradas como variáveis no modelo.

A Figura 2 apresenta as informações coletadas no perfil da @subwaybrasil, sendo que a seta aponta para o que representa um exemplo de postagem colaborativa, em que dois perfis fizeram uma parceria. Por meio da API do Instagram, tem-se informações de que essa publicação se trata de um vídeo, que tem menções e não tem *hashtags*, por exemplo.

Figura 2. Ilustração de informações coletadas em páginas do Instagram.




Fonte: Os autores (2023)

Nota: Em vermelho, destacou-se algumas informações coletadas em cada publicação. Disponível em: https://www.instagram.com/p/CWlNw5NLMoo/?img_index=1. Acesso em: 13/09/2023.

## 4.2 ANÁLISE

A análise de dados é realizada com o objetivo de entender como diferentes variáveis, como o número de curtidas em postagens, o conteúdo das legendas, as *hashtags* utilizadas e a frequência das postagens, afetam o nível de engajamento nas redes sociais das empresas de *Food Service*. Nesta etapa, fez-se inicialmente, uma análise sobre temas de interesse dos consumidores, indicando quais tópicos podem ser mais eficazes em atrair e envolver o público das empresas de alimentação. Essa análise foi realizada utilizando-se de técnicas de Processamento de Linguagem Natural (PLN) para analisar o conteúdo textual das legendas das

postagens, e *hashtags* associadas e para posterior extração de assuntos abordados em cada publicação. O PLN é um ramo da Inteligência Artificial que lida com a forma como as máquinas entendem e lidam com as linguagens humanas. De um modo geral, o texto coletado passou por dois processos: o Pré-processamento e a Extração de Palavras-Chave.

O Pré-processamento envolveu etapas como a remoção de caracteres especiais, pontuações e espaços em branco extras. Também foram realizadas a *tokenização* (divisão do texto em palavras ou *tokens*) e a conversão para letras minúsculas para garantir consistência. Já a Extração de Palavras-Chave, consistiu em uma análise de frequência lexical para extrair palavras-chave dos textos. Para realizar uma análise de frequência lexical correta, foi necessário considerar os devidos pré-processamentos. Visto que não é desejável que conjugações distintas de um mesmo verbo sejam contabilizadas como palavras diferentes. Da mesma forma, não se deseja que artigos como "o" e "a" constem na amostra, pois certamente constariam entre as palavras mais usadas sem que isso acrescentasse na análise. Assim, no pré-processamento, reduziu-se o tamanho do vocabulário e simplificou-se algumas formas lexicais, garantindo, assim, que o algoritmo consiga extrair informações relevantes e representativas de cada empresa.

Primeiramente, utilizou-se uma lista de *stopwords* em português disponível na biblioteca NLTK, para removermos essas palavras do texto. Essa lista contém artigos “o” e “uma”, preposições como “de” e “em”, entre outras palavras frequentes no idioma. À essa lista acrescentamos manualmente algumas formas contraídas de verbos ou preposições, como “tá” e “pra”. Em seguida, passou-se o texto por uma simplificação lexical com dois processos: *Stemização* e *Lematização*. O primeiro refere-se ao processo de reduzir as palavras flexionadas à sua raiz. O segundo é responsável por representar as palavras através do infinitivo para os verbos, e do masculino singular no caso de substantivos e adjetivos. Para esta tarefa, utilizou-se o lematizador da biblioteca *spaCy*. Enfim, avaliou-se a Frequência Lexical, que consiste na frequência que as palavras aparecem dentro do texto. Dessa forma, foram definidos os temas que apareciam com maior frequência nas postagens do segmento estudado.

Isso permitiu identificar os tópicos mais discutidos nas postagens, destacando quais termos eram mais relevantes. Após essa etapa, ainda houve um estudo exploratório dos resultados para que os principais temas ou tendências fossem evidenciadas.

Por fim, o tratamento de dados foi conduzido de modo a realizar:

- A eliminação dados inconsistentes: Foram identificados e excluídos dados inconsistentes, como número de *likes* oculto em certas postagens, o que poderia afetar a validade das análises.
- A codificação de Variáveis Categóricas: Quando as postagens incluíam informações categóricas, como *hashtags* ou categorias de conteúdo, foram aplicadas técnicas de codificação apropriadas, como a criação de variáveis *dummy*, para tornar esses dados adequados para análise estatística.

Na sequência, esses temas foram avaliados com base na relevância para o modelo considerando-se a pontuação por *Gain Ratio* (Ganho de Informação) proposto em QUINLAN (1993), que se trata de uma relação entre o ganho de informação e a informação intrínseca do atributo. Assim, foram definidos os temas mais relevantes ao consumidor:

- Promoções: Relacionado a posts que comunicam promoções, ofertas, descontos, prêmios, cupons, brindes, itens grátis, *cashback*.
- Bem-estar: Relacionado a informação de itens saudáveis, fitness, *light*.

- Embaixador: Relacionado a embaixadores da marca, que geralmente são influenciadores digitais escolhidos por sua relevância e capacidade de influenciar o público-alvo.
- Novidade: Relacionado ao lançamento de novos produtos.
- Duelo: Referente a posts com duelos ou batalha de produtos, votações.
- Segurança: Relacionado à informação de protocolos de segurança, higiene, cuidado e saúde.
- Eventos: Relacionado a eventos comemorativos e festivos, como aniversários, festivais, Natal, páscoa, Copa do Mundo e eventos esportivos, shows de TV e dias especiais.
- Limitada: Referente a edições limitadas de produtos.
- Delivery: Relacionado às informações sobre delivery de produtos.

Para exemplificar a classificação realizada, a Figura 3 apresenta uma postagem classificada como:

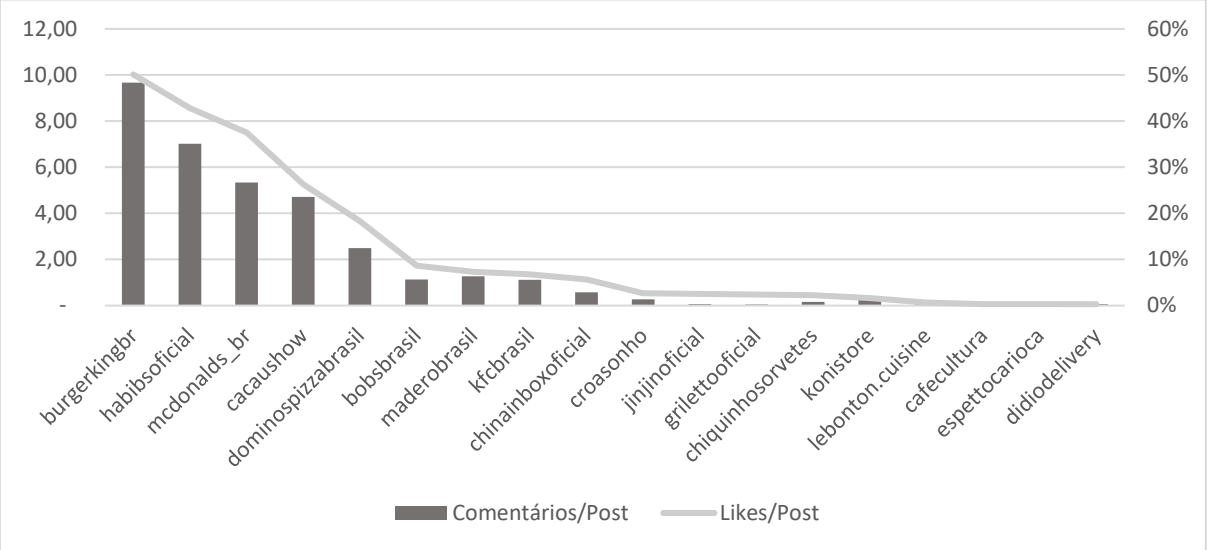
- Duelo: por se referir a uma batalha de produtos;
- Promoções: por relacionar-se a uma oferta de frete grátis;
- Eventos: por entrar na categoria de se referir a shows de TV, como o *Big Brother* Brasil (BBB); e
- *Delivery*: por oferecer essa opção.

Figura 3. Ilustração de publicação do Instagram.

Nota: Disponível em: https://www.instagram.com/p/Cn0GQuOP1LT/ Acesso em: 13/09/2023.
Fonte: Os autores (2023)

Uma vez que as publicações foram classificadas de acordo com seus temas, realizou-se uma análise descritiva dos dados. Considerando-se o número de *likes* por número de publicações, as empresas mais bem-sucedidas em envolver seu público-alvo, gerando mais interações e impacto em suas postagens foram a Burger King Brasil, Habib's e McDonald's Brasil, conforme pode-se visualizar no Gráfico 1.

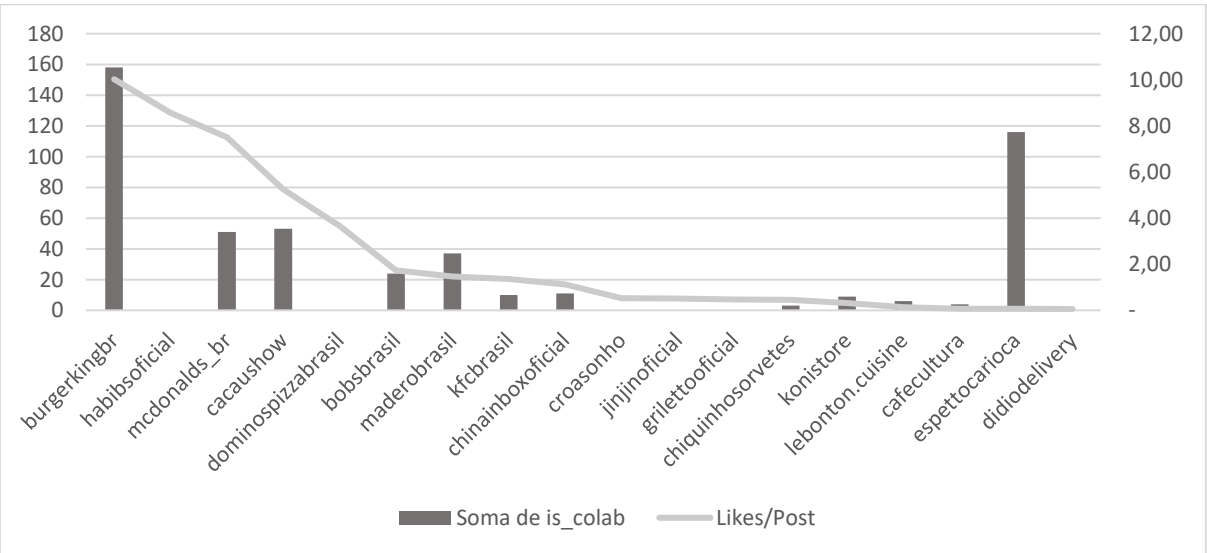
Gráfico 1. Número de *Likes* e Comentários por Número de Posts




Fonte: Os autores (2023)

Outro ponto importante a ser observado é que as variáveis, de um modo geral, não apresentam uma correlação clara com o número de *likes*. Isso fica evidente, por exemplo, ao analisar os posts colaborativos no Gráfico 2, onde se constata a existência de empresas que registram um maior volume de colaborações, contudo, obtendo-se um número reduzido de reações positivas (*likes*). Portanto, acredita-se que exista uma influência decorrente da interação complexa de diversas variáveis, tornando essencial o desenvolvimento de um modelo matemático que possa integrá-las de forma abrangente.

Gráfico 2. Número de *Likes* e Posts Colaborativos por Número de Posts




Fonte: Os autores (2023)

Após os resultados da análise, buscou-se identificar quais fatores contribuem para o aumento do engajamento nas redes sociais e estudar um modelo matemático capaz de auxiliar

as empresas a adaptarem suas estratégias com base nesses achados. Esse modelo incorporado ao simulador, tem potencial de orientar estratégias de postagens no Instagram de forma a otimizar o engajamento no setor de *Food Service.*

## 4.3 MODELAGEM MATEMÁTICA

Em seguida, realizou-se a modelagem dos dados para constituição do modelo matemático utilizando-se a linguagem de programação Python. Primeiramente, os dados a serem analisados foram divididos em conjuntos de treino (80%) e teste (20%). Por meio da análise de dados previamente coletados, ajustou-se o modelo para capturar as correlações entre essas variáveis e o nível de engajamento, permitindo-se prever como diferentes estratégias afetarão as interações com as postagens. Após o treinamento, avaliou-se o desempenho do modelo usando métricas de avaliação adequadas e ajustou-se o modelo alterando seus parâmetros de modo a otimizar as previsões.

Para a análise do modelo matemático, testou-se os seguintes algoritmos de aprendizado de máquina: *Random Forests*, *Support Vector Machines (SVM)*, *Neural Network* (redes neurais), *Gradient Boosting, AdaBoost e kNN (k-Nearest Neighbors)*. Em seguida, fez-se a escolha do modelo, com base em medidas de performance*:*

- O AUC (*Area Under the Curve*) avalia a capacidade discriminativa de um modelo de classificação. A curva ROC (*Receiver Operating Characteristic*) é plotada com a taxa de verdadeiros positivos no eixo y e a taxa de falsos positivos no eixo x. O AUC representa a área sob essa curva ROC. Quanto maior o valor do AUC (varia de 0 a 1), melhor o modelo em distinguir entre as classes.
- CA (*Accuracy*): A acurácia é uma métrica simples que mede a taxa de predições corretas feitas pelo modelo em relação ao total de predições.
- F1 Score: O F1 Score é calculado como a média harmônica entre *precision* (precisão) e *recall* (revocação). É particularmente útil quando há um desequilíbrio entre as classes do problema.
- *Precision* (Precisão): A precisão mede a proporção de previsões positivas corretas feitas pelo modelo em relação ao total de previsões positivas. Essa é uma métrica relevante em problemas onde é crítico minimizar os falsos positivos, ou seja, quando classificar algo como positivo quando é na verdade negativo é custoso ou prejudicial.
- *Recall* (Revocação): É uma métrica importante em problemas onde é crucial identificar todos os casos positivos, mesmo que isso resulte em alguns falsos positivos. Ou seja, *recall* pode ser usada em uma situação em que os Falsos Negativos são considerados mais prejudiciais que os Falsos Positivos.

Por fim, uma vez que o modelo foi considerado satisfatório, ele foi usado para fazer previsões sobre o número de *likes* em novas postagens do Instagram com base nas variáveis de entrada. Para que os resultados da análise e modelagem matemática fossem aplicados de forma prática e interativa pelas empresas, a etapa de criação do simulador foi crucial. O simulador permite aos usuários inserirem parâmetros específicos e prever o desempenho futuro das empresas no Instagram com base nas análises realizadas.

## 4.4 CRIAÇÃO DO SIMULADOR E DISPONIBILIZAÇÃO NA INTERNET

Em seguida, o simulador foi disponibilizado na internet por meio de uma plataforma *web* acessível ao público. Isso permitiu que outras pessoas interessadas pudessem utilizar a ferramenta e explorar as previsões com base nos dados disponíveis. Para essa etapa, uma

interface de usuário foi projetada e desenvolvida para o simulador. Ela permitiu aos usuários inserirem os parâmetros desejados de acordo com suas necessidades e objetivos. A interface foi projetada de forma intuitiva e amigável para facilitar a interação.

A disponibilização do simulador na internet foi fundamental para tornar as análises e previsões acessíveis a um público mais amplo. O simulador foi hospedado em um servidor *web* confiável e escalável, a parte de *backend* está disponível na nuvem do Google e o *frontend* disponível na nuvem da Microsoft. Isso garantiu que ele estivesse sempre disponível para acesso *online* e pudesse lidar com mais usuários simultaneamente. Foram implementadas medidas de segurança para proteger os dados inseridos pelos usuários no simulador. O simulador foi disponibilizado publicamente na internet, permitindo que pesquisadores, profissionais de *marketing*, e interessados pudessem testá-lo gratuitamente através do link https://simulacom.com.br/food bastando que seja feito um cadastro inicial.

## 5 ANÁLISE DOS RESULTADOS

Nesta seção, apresenta-se os passos que se seguiram após a coleta de dados. Apresentando-se a análise de dados, as etapas da criação do modelo e a criação do simulador e sua disponibilização na internet.

### 5.1 Modelagem matemática

Os testes iniciais foram conduzidos utilizando o número de "likes" como parâmetro de avaliação, adotando, consequentemente, um modelo de regressão. No entanto, os resultados obtidos com base nas métricas de desempenho se revelaram insatisfatórios. Diante disso, a abordagem foi alterada para transformar a variável de resposta em uma categorização baseada em faixas de números de "likes," efetivando assim um modelo de classificação. A determinação das faixas foi realizada com base em uma subdivisão em três partes ou percentis. Os resultados obtidos para os dados de Treino e Teste deste modelo de classificação, estão descritos na Tabela 1 e Tabela 2.

Tabela 1. Scores para dados de Treino

| ***Model*** | ***AUC*** | ***CA*** | ***F1*** | ***Precision*** | ***Recall*** |
|---|---|---|---|---|---|
| ***Random Forest*** | 0,886 | 0,712 | 0,706 | 0,711 | 0,712 |
| ***Gradient Boosting*** | 0,845 | 0,657 | 0,639 | 0,652 | 0,657 |
| ***AdaBoost*** | 0,837 | 0,668 | 0,664 | 0,669 | 0,668 |
| ***kNN*** | 0,835 | 0,661 | 0,662 | 0,670 | 0,661 |
| ***Neural Network*** | 0,825 | 0,639 | 0,637 | 0,645 | 0,639 |
| ***SVM*** | 0,350 | 0,487 | 0,461 | 0,482 | 0,487 |

Fonte: Os autores (2023)

Tabela 2. Scores para dados de Teste

| ***Model*** | ***AUC*** | ***CA*** | ***F1*** | ***Precision*** | ***Recall*** |
|---|---|---|---|---|---|
| ***Gradient Boosting*** | 0,838 | 0,649 | 0,630 | 0,643 | 0,649 |
| ***Random Forest*** | 0,833 | 0,657 | 0,649 | 0,651 | 0,657 |

| | | | | | |
|---|---|---|---|---|---|
| ***kNN*** | 0,799 | 0,624 | 0,625 | 0,634 | 0,624 |
| ***Neural Network*** | 0,794 | 0,606 | 0,602 | 0,610 | 0,606 |
| ***AdaBoost*** | 0,765 | 0,605 | 0,601 | 0,604 | 0,605 |
| ***SVM*** | 0,468 | 0,476 | 0,451 | 0,470 | 0,476 |

Fonte: Os autores (2023)

Para a seleção do modelo foram avaliadas algumas métricas como o AUC, CA, F1, *Precision* e *Recall*, que são amplamente utilizadas para avaliar o desempenho de modelos de classificação em problemas de aprendizado de máquina e mineração de dados. Cada uma dessas métricas fornece uma perspectiva diferente sobre a qualidade e a eficácia de um modelo de classificação. A seguir analisa-se as medidas de performance com base no modelo do *Gradient Boosting:*

- O AUC (*Area Under the Curve*): Considerando-se os dados de teste, o modelo pode ser considerado "bom" como classificador, visto que seu valor foi acima de 80%, e seria considerado "excelente" caso excedesse os 90% (LAURETTO, 2023).
- CA (*Accuracy*): O modelo apresentou 64,9% de previsões corretas em relação ao total.
- F1 Score: O F1 Score foi de 63,0% representando a capacidade do modelo em identificar positivos verdadeiros e minimizar falsos positivos.
- *Precision* (Precisão): O modelo apresentou 64,3% de precisão.
- *Recall* (Revocação) de 64,9%, significando dentre todas as situações de classe Positivo como valor esperado, o percentual de quantas estão corretas.

Na definição das métricas, é importante avaliar qual o prejuízo em se obter falsos positivos ou negativos. No primeiro caso, um Falso Positivo seria o caso de uma postagem ser de alto engajamento, porém ela não ser reconhecida pelo modelo. Isso poderia representar oportunidades perdidas, por deixar de postar algo que seria relevante. Já os Falsos Negativos, ocorreriam caso o modelo previsse que uma postagem teria sucesso, porém ela não apresentasse bons índices de engajamento. O prejuízo neste caso, seria em termos de sucesso de engajamento a longo prazo que não seria sustentado, uma vez que os seguidores passariam a perceber que a postagem não é realmente relevante para eles. Considerou-se que ambos os casos são indesejáveis.

Para se evitar a seleção de modelos sobreajustados (*overfitting*) aos dados de treinamento e de baixa capacidade de generalização, fez-se o treinamento do modelo por meio da aplicação do método de validação cruzada (*Cross Validation***)** *k-fold*. Segundo Shalev-Shwartz e Ben-David (2014), o método *k-fold* consiste em particionar a base de dados aleatoriamente em k subconjuntos de dimensões similares e, para cada configuração do algoritmo, realizar *k* treinamentos, separando um subconjunto por vez para validação. Desse modo, todas as instâncias são utilizadas *k-1* vezes para treinamento e uma vez para validação. A medida de performance de cada configuração é obtida pela média do erro quadrático médio das *k* validações. Utilizou-se neste estudo, o método *5-fold,* e os resultados das métricas avaliados estão descritos na Tabela 3.

Tabela 3. *Scores* para a Validação Cruzada com 5 *folds*

| ***Model*** | ***AUC*** | ***CA*** | ***F1*** | ***Precision*** | ***Recall*** |
|---|---|---|---|---|---|
| ***Gradient Boosting*** | 0,829 | 0,643 | 0,624 | 0,636 | 0,643 |

| | | | | | |
|---|---|---|---|---|---|
| ***Random Forest*** | 0,827 | 0,651 | 0,642 | 0,644 | 0,651 |
| ***kNN*** | 0,793 | 0,623 | 0,623 | 0,630 | 0,623 |
| ***Neural Network*** | 0,792 | 0,607 | 0,598 | 0,603 | 0,607 |
| ***AdaBoost*** | 0,775 | 0,635 | 0,628 | 0,629 | 0,635 |
| ***SVM*** | 0,556 | 0,556 | 0,357 | 0,412 | 0,396 |

Fonte: Os autores (2023)

O modelo selecionado, portanto, foi o *Gradient Boosting* que consiste em um algoritmo de aprendizado de máquina que se destaca na criação de modelos preditivos precisos a partir de dados complexos. Ele é particularmente eficaz para tarefas de classificação, como neste caso de se prever uma faixa de número de *likes* com base em várias variáveis. O modelo apresentou um AUC de 0,829, e métricas de F1 Score, a precisão e o recall estão acima de 60%.

Após a concepção do modelo e sua validação, cumpre-se a sua função de auxiliar as empresas de alimentação na tomada de decisões em relação às estratégias de mídia social. As empresas podem utilizar o simulador para experimentar diferentes cenários, ajustar suas abordagens e antecipar os resultados esperados de suas postagens no Instagram.

## 5.2 Criação do simulador

Para a criação do simulador utilizou-se linguagens de programação Python, e ferramentas específicas de desenvolvimento *web* para garantir que o simulador seja uma ferramenta intuitiva e de fácil utilização. O projeto de construção do simulador, utilizou metodologias ágeis e de *growth hacking* para garantir que os esforços estejam orientados para a demanda de empresas do segmento. A Figura 4 apresenta a ilustração da interface do simulador.

Figura 4. Ilustração da tela do simulador

Fonte: Os autores (2023)

Quanto à sua disponibilização na internet, desenvolveu-se uma interface de usuário amigável e de uma programação *web* responsiva. Para o desenvolvimento da Infraestrutura (*Back end*) construiu-se um *Webservice* para realizar o processamento dos dados e execução de modelos de inteligência artificial. Já para o desenvolvimento da Interface do usuário (*Front end*): considerou-se desde a fluidez da usabilidade para as empresas (*UX design*) até a segurança e Lei Geral de Proteção de Dados (LGPD). Este estudo respeitou todas as políticas de privacidade e diretrizes éticas da coleta de dados do Instagram.

A disponibilização *online* do simulador permite que empresas do setor de *Food Service*, independentemente de sua localização geográfica, tenham acesso fácil e imediato a essa ferramenta valiosa. Essa acessibilidade visa a facilitar a integração do simulador nas estratégias de mídia social das empresas, tornando-o acessível a qualquer momento e de qualquer lugar.

## 6 CONCLUSÃO /CONTRIBUIÇÃO

Ao longo deste trabalho, apresentou-se uma solução a um desafio crítico enfrentado por empresas de alimentação: a otimização das estratégias de postagens no Instagram para aumentar o engajamento e o alcance do público-alvo. Através de uma abordagem multidisciplinar que incorporou princípios de empreendedorismo, inovação e tecnologia, desenvolvemos um simulador que oferece às empresas ferramentas práticas e personalizadas para aprimorar suas estratégias de mídia social.

A execução do modelo matemático representou um avanço significativo, permitindo a previsão do impacto de diferentes variáveis nas postagens no Instagram. Além disso, a criação e implementação de um simulador de postagens no Instagram representa um avanço significativo na busca por estratégias eficazes de *marketing* digital para maximizar o número

de curtidas nas publicações. Além disso, a disponibilização do simulador na internet tornou esta ferramenta acessível a empresas de todos os tamanhos e localizações geográficas. Isso não apenas democratiza o acesso à inovação e tecnologia, mas também capacita empresas de *Food Service* a e otimizarem suas estratégias de mídia social, impulsionando seu sucesso digital.

A contribuição deste simulador vai além da simples otimização das estratégias de postagens; ele preenche uma lacuna notável ao fornecer uma ferramenta específica e prática para um setor que carece de recursos especializados. Além de ajudar as empresas a aumentar sua visibilidade e atratividade *online*, essa ferramenta também promove a eficiência no uso dos recursos de *marketing*, tornando-se um valioso aliado na busca por resultados concretos.

Como potencial de estudos futuros, sugere-se estudos sobre a alimentação de informações em tempo real ao simulador. Isso representaria um passo significativo em direção a uma abordagem mais dinâmica e adaptativa para o *marketing* de alimentos em redes sociais. Outra área promissora para estudos futuros envolve a expansão da análise para outras redes sociais, como Twitter e TikTok, além do Instagram. Cada uma dessas plataformas possui características únicas e públicos diferentes, o que poderia fornecer insights valiosos para as estratégias de *marketing* digital no setor de *Food Service*.

## AGRADECIMENTOS

Os autores agradecem ao CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico e à FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo pelo apoio fornecido ao longo deste trabalho.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABRASEL. Bares e restaurantes estão mais digitais na gestão e mais presentes nas redes sociais, diz pesquisa exclusiva da Abrasel. 30/05/2023a. Disponível em: https://abrasel.com.br/noticias/noticias/bares-e-restaurantes-estao-mais-digitais-na-gestao-e-mais-presentes-nas-redes-sociais-diz-pesquisa-exclusiva-da-abrasel/. Acesso em: 07/09/2023.

ABRASEL. Franquias registram crescimento e impulsionam retomada do setor. 07/07/2023b. Disponível em: https://abrasel.com.br/revista/mercado/franquias-de-alimentacao-registram-crescimento-expressivo-e-impulsionam-retomada-do-setor/. Acesso em: 07/09/2023.

ABRASEL. Pesquisa CREST®: Foodservice brasileiro segue tendência de crescimento. 08/07/2019. Disponível em: https://ba.abrasel.com.br/noticias/noticias/pesquisa-crest-foodservice-brasileiro-segue-tendencia-de-crescimento/. Acesso em: 06/09/2023.

ASSOCIAÇÃO BRASILEIRA DE AGÊNCIAS DE PUBLICIDADE (ABAP). **O Impacto da Publicidade na Economia do Brasil,** 2015. Disponível em: https://www.abap.com.br/wp-content/uploads/2021/07/impactodapublicidadenaeconomiadobrasil.pdf. Acesso em: 20/06/2023.

BAG, S., SRIVASTAVA, G., BASHIR, MMA, KUMARI, S., GIANNAKIS, M. E CHOWDHURY, AH. Jornada dos clientes nesta era digital: entendendo o papel das tecnologias de inteligência artificial no envolvimento do usuário e conversão, Benchmarking: An International Journal, Vol. 29 No. 7, 2022. Disponível em: https://doi.org/10.1108/BIJ-07-2021-0415 Acesso: 01/09/2023.

CHAIT, G. "Half the money I spend on advertising is wasted; the trouble is I don't know which half.", **B2B Marketing,** 18 mar. 2015. Disponível em: https://www.b2bmarketing.net/en-gb/resources/blog/half-money-i-spend-advertising-wasted-trouble-i-dont-know-which-half. Acesso em: 18/07/2023.

CRESSWELL, Kathrin et al. Understanding public perceptions of COVID-19 contact tracing apps: Artificial intelligence–enabled social media analysis. **Journal of medical Internet research**, v. 23, n. 5, p. e26618, 2021.

D'ANGELO, B. J., Social media community management: implications for business communication curriculum, in Proceedings of the 75th annual convention of the Association for Business Communication, pp. 27-30, 2010.

DOSSENA, C. et al. Restaurants and social media: rethinking organizational capabilities and individual competencies. **Journal of tourism Futures**, v. 7, n. 1, p. 20-39, 2021.

FERNANDEZ-LUQUE, Luis; IMRAN, Muhammad. Humanitarian health computing using artificial intelligence and social media: A narrative literature review. **International journal of medical informatics**, v. 114, p. 136-142, 2018.

FERNANDEZ-LUQUE, Luis; IMRAN, Muhammad. Humanitarian health computing using artificial intelligence and social media: A narrative literature review. International journal of medical informatics, v. 114, p. 136-142, 2018.

FOU, Augustine. In Digital, 'Wanamaker's 50%' Is Known. It's Also Worse Than That. **Forbes,** 19 dez. 2020. Disponível em: https://www.forbes.com/sites/augustinefou/2020/12/19/in-digital-wanamakers-50-is-known-its-also-worse-than-that/?sh=5d9ed9d8714b. Acesso em: 18/07/2023.

FREITAS, B. V. R.; SHIGAKI, H. B. A Experiência do Cliente no Setor de *Food Service*: Análise da Literatura e Proposição de uma Agenda de Pesquisa. ***Marketing* & Tourism Review**, 2022.

GASPARINI, A. A influência das redes sociais no *Food Service*. 12/07/2023. Disponível em: https://redefoodservice.com.br/2023/07/a-influencia-das-redes-sociais-no-food-service/ Acesso em: 07/09/2023.

H. SARMIENTO, "How artificial intelligence can benefit the social media user," May 2020, https://medium.com/clyste/how-artificialintelligence-can-benefit-the-social-media-useraeaefd24e0a7 Acesso:10/08/2023.

KANTAR IBOPE MÍDIA. **Inside advertising**. 2021. Disponível em: https://www.kantaribopemedia.com/inside-advertising-download/. Acesso em: 18 jul. 2021.

KAUFMAN, Dora; SANTAELLA, Lucia. O papel dos algoritmos de inteligência artificial nas redes sociais. Revista Famecos, v. 27, n. 1, p. e34074-e34074, 2020.

KUKANJA, M.; OMERZEL, D. G.; KODRIČ, B. Ensuring restaurant quality and guests' loyalty: an integrative model based on *marketing* (7P) approach. Total Quality Management & Business Excellence, v. 28, n. 13–14, p. 1509–1525, 10 nov. 2017.

LAURETTO, M. S. Introduction to Supervised Learning Performance Evaluation. Disponível em: https://www.each.usp.br/lauretto/supervised_learning_intro/parte03.pdf Acesso em: 10/09/2023.

MCCLENEGHAN, J.S., "PR executives rank 11 communication skills", Public Relations Quarterly, Vol. 51 No. 4, pp. 42-46, 2006.

OWUSU, Priscilla N. et al. Artificial intelligence applications in social media for depression screening: A systematic review protocol for content validity processes. **Plos one**, v. 16, n. 11, p. e0259499, 2021.

QUINLAN, J. R. C4.5: programs for machine learning. Morgan Kaufmann Publishers, 1993.

ROY, A.; DIONNE, C., "Social media and SMEs: how to remove the barriers and improved organisational performance?", in proceedings of 3rd European Conference on Social M di R h Media ResearchEM Normandie, Caen, France, pp. 360-368, 2016.

S. GREENGARD, “What is artificial intelligence?” May 2019, https://www.datamation.com/artificialintelligence/what-is-artificial-intelligence.html Acesso: 20/06/2023.

SADIKU, M. N., ASHAOLU, T. J., AJAYI-MAJEBI, A., MUSA, S. M. Artificial intelligence in social media. **International Journal of Scientific Advances**, v. 2, n. 1, p. 15-20, 2021.

SBVC. Ranking 300 Maiores Empresas do Varejo Brasileiro 2022 SBVC – SBVC – Sociedade Brasileira de Varejo e Consumo. São Paulo, 2022. Disponível em: https://sbvc.com.br/ranking-300-maiores-empresas-do-varejo-brasileiro-2022-sbvc/. Acesso em: 04/09/2023.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. Understanding Machine Learning: From Theory to Algorithms. New York: Cambridge University Press, 1. ed., 2014.

State of Inbound *Marketing* Trends 2022. Disponível em: https://www.hubspot.com/hubfs/2022_State-of-Inbound-*Marketing*-Trends_V712.pdf Acesso em: 02/09/2023.

TENCH, R.; MORENO, A. Mapping communication management competencies for European practitioners: ECOPSI an EU study, Journal of Communication Management, Vol. 19 No. 1, pp. 39-61, 2015.

THINKBOX. **Profit Ability: The Business Case for Advertising**. Disponível em: https://www.thinkbox.tv/research/thinkbox-research/profit-ability-the-business-case-for-advertising/. Acesso em: 20/06/023.

TUSSYADIAH, I.P.; FESENMAIER, D.R., Mediating tourist experiences: access to places via shared videos, Annals of Tourism Research, Vol. 36 No. 1, pp. 24-40, 2009.