



APRENDIZADO DE MÁQUINA APLICADO À ANÁLISE DE SENTIMENTOS NO TURISMO: REVISÃO SISTEMÁTICA E AGENDA DE PESQUISA

MACHINE LEARNING APPLIED TO SENTIMENT ANALYSIS IN TOURISM: A SYSTEMATIC REVIEW AND FUTURE RESEARCH AGENDA

APRENDIZAJE AUTOMÁTICO APLICADO AL ANÁLISIS DE SENTIMIENTOS EN EL TURISMO: UNA REVISIÓN SISTEMÁTICA Y AGENDA DE INVESTIGACIÓN FUTURA

Bruno Homann Zilli — UFPR¹
Melise de Lima Pereira — UFPR²

ISSN 2594-8497



Licenciada por Creative
Commons 4.0 / Internacional
CC BY 4.0

Submetido em: 26/12/2025
Aprovado em: 16/03/2026
Avaliado em pares
Editor: Izac Bonfim

RESUMO

O turismo é um fenômeno baseado em informação. A análise de dados é fundamental para o setor. O Aprendizado de Máquina (ML) e a Análise de Sentimentos (AS) oferecem uma solução para a automatização da análise de dados de UGC (conteúdo gerado por usuário). No Brasil, entretanto, a complexidade desses modelos ainda limita sua aplicação metodológica e prática. A pesquisa justifica-se na crescente ascensão do uso do UGC e Inteligência Artificial no turismo. O objetivo é elaborar uma agenda de pesquisa futura da aplicação de ML para a pesquisa com AS com dados textuais de UGC em turismo. Os objetivos específicos são: i) compreender a utilização do ML nos estudos em turismo e ii) identificar possíveis lacunas nas pesquisas sobre o uso de ML no turismo. Para isso foi realizada uma revisão sistemática de literatura nas bases de periódicos científicos Scopus e Web of Science. Foram escolhidos 10 artigos para revisão a partir de ranqueamento com o Methodi Ordinatio. Os resultados mostram a predominância de dois tipos de estudo: aquele cujo foco está em comparar e/ou desenvolver algoritmos para pesquisa, e aquele que aplica o ML para AS como uma de suas ferramentas de pesquisa. Os estudos também abordam questões teóricas, destacando-se a Computação Afetiva como eixo comum. Os principais modelos de ML são algoritmos supervisionados ou semi-supervisionados, com aplicação baseada em Naïve Bayes ou SVM e utilização de Topic Modelling ou ASBA para mineração dos sentimentos no texto. Estudos futuros deverão desenvolver modelos mais complexos e contextualmente adaptados, além de aprofundar a relação teórica entre turismo e Computação Afetiva. Ademais, é fundamental que os avanços obtidos sejam disseminados, tanto para os profissionais do setor quanto para os turistas, de modo a potencializar sua aplicação prática e promover um turismo mais inteligente e centrado na experiência humana.

Palavras-chave: Turismo; UGC; Inteligência Artificial; Revisão Sistemática; Agenda de pesquisa.

ABSTRACT

Tourism is an information-based phenomenon. Data analysis is fundamental to the sector. Machine Learning (ML) and Sentiment Analysis (SA) offer a solution for automating the analysis of user-generated content (UGC) data. In Brazil, however, the complexity of these models still limits their methodological and practical application. The relevance of this research lies in the growing use of UGC and artificial intelligence in tourism. The objective is to develop a future research agenda for the application of ML to sentiment analysis research using textual UGC data in tourism. The specific objectives are: (i) to understand the use of ML in tourism studies and (ii) to identify possible research gaps regarding the use of ML in tourism. To this end, a systematic literature review was conducted using the Scopus and Web of Science databases. Ten articles were selected for review based on a ranking using the Methodi Ordinatio. The results indicate the predominance of two types of studies: those focused on comparing and/or developing algorithms for research purposes, and those that apply ML to SA as one of their research tools. The studies also address theoretical issues, with Affective Computing standing out as a common axis. The main ML models are supervised or semi-supervised algorithms, predominantly based on Naïve Bayes or SVM, with the use of Topic Modelling or Aspect-Based Sentiment Analysis (ABSA) for sentiment mining from text. Future studies should develop more complex and contextually adapted models, as well as deepen the theoretical relationship between tourism and Affective Computing. Furthermore, it is essential that the advances achieved be disseminated both to industry professionals and to tourists, in order to enhance their practical application and promote smarter, more human experience-centered tourism.

Keywords: *Tourism; UGC; Artificial Intelligence; Systematic Literature Review; Future Research Agenda.*

RESUMEN

El turismo es un fenómeno basado en la información. El análisis de datos es fundamental para el sector. El aprendizaje automático (ML) y el análisis de sentimientos (AS) ofrecen una solución para automatizar el análisis de datos de contenido generado por el usuario (UGC). En Brasil, sin embargo, la complejidad de estos modelos aún limita su aplicación metodológica y práctica. Esta investigación se justifica por el creciente uso de UGC e inteligencia artificial en el turismo. El objetivo es desarrollar una agenda de investigación futura para la aplicación de ML a la investigación de AS con datos textuales de UGC en turismo. Los objetivos específicos son: i) comprender el uso de ML en estudios de turismo y ii) identificar posibles brechas en la investigación sobre el uso de ML en turismo. Para ello, se realizó una revisión sistemática de la literatura en las bases de datos de revistas científicas Scopus y Web of Science. Se seleccionaron diez artículos para la revisión con base en la clasificación utilizando Methodi Ordinatio. Los resultados muestran el predominio de dos tipos de estudios: aquellos enfocados en comparar y/o desarrollar algoritmos para la investigación, y aquellos que aplican ML a AS como una de sus herramientas de investigación. Los estudios también abordan cuestiones teóricas, destacando la Computación Afectiva como eje común. Los principales modelos de aprendizaje automático son algoritmos supervisados o semisupervisados, con aplicaciones basadas en Naïve Bayes o SVM y el uso de Topic Modelling o ASBA para el análisis de sentimientos en texto. Futuros estudios deberían desarrollar modelos más complejos y adaptados al contexto, así como profundizar en la relación teórica entre turismo y Computación Afectiva. Además, es fundamental que los avances obtenidos se difundan, tanto entre los profesionales del sector como entre los turistas, para potenciar su aplicación práctica y promover un turismo más inteligente centrado en la experiencia humana.

Palabras clave: *Turismo; UGC; Inteligencia Artificial; Revisión Sistemática; Agenda de Investigación.*

Como citar (APA 7ª ed.): Zilli, B. H., & Pereira, M. de L. (2026). Aprendizado de máquina aplicado à análise de sentimentos no turismo: revisão sistemática e agenda de pesquisa. *Ateliê do Turismo*, 10(1) 16 – 36, e24785. <https://doi.org/10.55028/a7r9kf38>

INTRODUÇÃO

A informação é o elemento fundamental do Turismo, sem o qual o planejamento estratégico do setor se torna inviável (Silva, 2023; Song et al., 2023). Por isso, a análise de dados é uma etapa essencial para o planejamento turístico (Neratika, 2016; Cruz, 2018; Gonçalves & Oliveira, 2023). Até décadas recentes, as dificuldades e entraves nos meios tradicionais de coleta e processamento de dados representavam um grande desafio para o Turismo. Gestores tinham que lidar com constantes limitações na disponibilidade e disseminação da informação (Romele et al., 2018; Porto et al., 2023; Santos et al., 2024).

As transformações tecnológicas recentes mudaram a forma como a informação é produzida e utilizada (Gursoy et al., 2024). O conteúdo gerado por usuários (UGC – *user generated content*) em plataformas de mídias sociais trouxe um novo paradigma para a pesquisa em turismo, mas também um novo desafio: as metodologias e técnicas de pesquisa clássicas não são capazes de lidar com o volume e a variedade de dados produzidos nas plataformas online (Silva, 2023; Maia et al., 2024).

Uma solução para esse desafio é a aplicação do Aprendizado de Máquina (ML- *machine learning*), um tipo de Inteligência Artificial (IA), no processo de análise de dados (Zang et al., 2024). A IA automatiza o processamento de dados e, combinada ao UGC, impulsiona a inovação no turismo (Delipetrev et al., 2020; Gonçalves & Oliveira, 2023; Porto et al., 2023). O uso de ML tem crescido no setor, sendo especialmente aplicado para pesquisas de previsão de demanda turística e de Análise de Sentimentos (AS), técnica computacional que classifica de forma automatizada as emoções expressas em textos (Alaei et al., 2017; Manosso & Ruiz, 2021; Xu et al., 2022; Sari et al., 2022). No Brasil, entretanto, a complexidade desses modelos ainda limita sua aplicação metodológica e prática (Gonçalves & Oliveira, 2023; Santos et al., 2024).

Assim, a presente pesquisa busca realizar uma revisão sistemática de literatura para responder a seguinte pergunta: quais os principais modelos de Aprendizado de Máquina utilizados no turismo para Análise de Sentimentos a partir de dados textuais de UGC? O objetivo principal é elaborar uma agenda de pesquisa futura da aplicação de ML para a pesquisa com AS com dados textuais de UGC em turismo. Os objetivos específicos da pesquisa são: i) compreender a utilização do Aprendizado de Máquina nos estudos em turismo e ii) identificar possíveis lacunas nas pesquisas sobre o uso de Aprendizado de Máquina no turismo.

Dentro desse escopo, pretende-se responder as seguintes questões:

P1: quais as vantagens e desvantagens dos principais modelos de ML utilizados para AS no turismo?

P2: quais as metodologias empregadas nas principais pesquisas sobre o tema?

P3: quais os principais campos de pesquisa inseridos nesse tema?

P4: quais as tendências do tema e como elas se relacionam com outros campos e temas de pesquisa?

A pesquisa justifica-se na crescente ascensão do uso do UGC e IA no turismo, seja no campo acadêmico ou em empresas e destinos turísticos. Metodologias de pesquisas quantitativas tradicionais muitas vezes não atingem o número mínimo de amostra, sofrem com vieses amostrais e/ou têm limitações orçamentárias e de escopo (Romele et al., 2018; Porto et al., 2023; Santos et al., 2024). Nesse sentido, o uso de IA para análise de UGC permite uma

abordagem longitudinal nas pesquisas quantitativas a partir de dados secundários. Além disso, embora o ML tenha potencial para a análise de dados no turismo, sua aplicação em pesquisas de turismo é recente e limitada (Lima et al., 2022). A seguir, a pesquisa apresenta a revisão de literatura; a metodologia; a análise dos resultados; a discussão dos resultados; as conclusões e as referências

REFERENCIAL TEORICO

Os textos online e outras formas de UGC transformaram a forma como a informação e a opinião são compartilhadas (Li et al., 2018; Mountasser et al., 2020; Silva, 2023). A análise de textos online assume um papel importante nos estudos de opinião e percepção em variadas áreas, visto que dados de UGC refletem de maneira eficiente os sentimentos e as opiniões dos usuários sobre produtos ou serviços (Nieves-Pavón et al., 2024). O turismo é entendido como um fenômeno fundamentado na informação, o que qualifica dados como a matéria-prima do planejamento turístico. Essa perspectiva coloca a emergência do UGC não só como fonte de dados estratégica, mas também como catalisadora da inovação no turismo (Li et al., 2018; Mountasser et al., 2020; Silva, 2023).

Cada vez mais a competitividade turística depende da habilidade dos gestores de extrair valor de dados digitais (Silva, 2023; Troisi et al., 2023). Pesquisadores e gestores têm se debruçado sobre o conteúdo gerado em microblogs, como o X (Chen et al., 2022), e avaliações em sites de viagens, como o TripAdvisor (Lyu et al., 2022). Essas plataformas de compartilhamento de opinião se tornaram centrais na experiência turística (Romele et al., 2018; Li et al., 2019; Manosso & Ruiz, 2021; Silva, 2023). Contudo, um dos desafios do UGC é sua natureza desestruturada, o que dificulta sua utilização para pesquisas ou monetização (Barbera et al., 2023; Maia et al., 2024). Nesse contexto, a IA destaca-se como uma das ferramentas mais eficazes para análise do UGC, pois permite automatizar os processos de coleta, monitoramento e estruturação de dados (Santos et al., 2024; Nnachi et al., 2024).

A IA pode ser definida como o campo da computação que estuda a habilidade de uma máquina aprender à semelhança de um ser humano, ou mesmo como o estudo da inteligência demonstrada por máquinas (Schank, 1991; Delipetrev et al., 2020). Essa inteligência se concretiza, na maior parte dos casos, na forma da execução de automática de tarefas que tradicionalmente dependeriam da cognição humana (Santos et al., 2024; Wu et al., 2025). O funcionamento de modelos de IA depende da disponibilidade de dados em grande volume e em grande variedade, por isso, por muitos anos a principal limitação para o avanço científico e comercial da IA foi a escassez de dados. Contudo, esse cenário mudou com a proliferação do Big Data e do UGC (Delipetrev et al., 2020; Porto et al., 2023; Santos et al., 2024; Wu et al., 2025).

O ML é um subcampo da IA (Santos et al. 2024) caracterizado por ser um modelo computacional capaz de aprender a partir da experiência, sem a necessidade de instruções explícitas de um ser humano (Herath & Mittal, 2022; Nnachi et al., 2024). O funcionamento de qualquer IA é intrinsecamente ligado à sua capacidade de aprendizado, e algoritmos de ML são projetados para aprender de maneira autônoma, utilizando grandes conjuntos de dados para identificar e reconhecer padrões - os quais são a base sobre a qual os modelos ajustam seu comportamento futuro e tomam decisões (Schank, 1991; Santos et al., 2024). Uma amostra de dados, chamada de "dados de treinamento", é utilizada para ensinar o algoritmo. Durante o treinamento, o algoritmo analisa os dados, prevê e atribui classificações ou valores e identifica padrões presentes na amostra. Após esse processo, o modelo é capaz de reconhecer os mesmos

padrões em novas amostras de dados, aplicando o conhecimento adquirido (Herath & Mittal, 2022; Leelawat et al., 2022).

Há três tipos principais de treinamento de ML (Delipetrev et al., 2020; Nnachi et al., 2024): (1) supervisionado, com dados estruturados, em que o modelo aprende a reconhecer padrões específicos; (2) não supervisionado, com dados não estruturados, em que os padrões são descobertos sem orientação; e (3) semi-supervisionado, que combina ambos os tipos de dados. Essa capacidade torna o ML especialmente eficaz na análise de UGC não estruturado, automatizando classificações e atribuições. Aplicações recentes incluem a previsão de sentimentos em avaliações de hotéis (Marigliano, 2024), a mensuração da experiência de marca do destino (Calderón-Fajardo et al., 2024) e a previsão de demanda turística (Hu et al., 2025).

Um dos métodos mais beneficiados pela aplicação do ML é a AS, método para extração de informação subjetiva (sentimentos e emoções) de dados textuais (Kumar, 2021; Manosso & Ruiz, 2021; Sari et al., 2022). A AS transforma esse conteúdo em dados quantificáveis por meio da classificação dos textos em sentimentos positivos, negativos e neutros. Embora essa classificação possa ser feita manualmente, trata-se de um processo demorado e limitado em escala (Xu et al., 2022; Weismayer et al., 2021). Para automatizá-lo, utilizam-se léxicos ou algoritmos de ML. Léxicos atribuem valores pré-definidos a palavras, sendo simples de aplicar, porém menos sensíveis a nuances temáticas (Kumar, 2021; Xu et al., 2022; Weismayer et al., 2021). Já os algoritmos de ML são treinados com base nos próprios dados, oferecendo maior precisão, sobretudo em contextos específicos (Nnachi et al., 2024).

Estudos recentes apontam resultados melhores com ML do que com léxicos. Liu et al. (2023), por exemplo, identificaram maior acurácia na detecção de emoções em avaliações de restaurantes, enquanto Lin et al. (2024) utilizaram AS para analisar o papel do valor emocional em experiências de turismo na natureza. Assim, juntos, ML e AS formam um método eficaz e confiável para estratégias baseadas em dados textuais no planejamento turístico. A seguir, são apresentados a metodologia e os procedimentos de pesquisa do estudo.

METODOLOGIA

Esta pesquisa é uma revisão sistemática com o intuito de analisar a evolução e as tendências do campo de estudo e responder as perguntas de pesquisa propostas. A pesquisa se caracteriza pelo uso de métodos quantitativos e de dados secundários (Veal, 2011). As bases de periódicos selecionadas para a pesquisa foram a Scopus e a Web of Science (WoS), pois são algumas das bases de periódicos científicos com maior produção em estudos de turismo (Gursoy et al., 2024; Maia et al., 2024). O procedimento de pesquisa é ilustrado na Figura 1 a seguir. Esse método de análise permite diminuir o viés da pesquisa ao estabelecer critérios claros para análise da informação (Aria & Cuccurullo, 2017; Li et al., 2019; Manosso & Ruiz, 2021; Gursoy et al., 2024; Maia et al., 2024).

Fig. 1



Procedimentos metodológicos

Nota: Fonte: elaboração própria (2026).

O procedimento de pesquisa seguiu da seguinte maneira:

1. Pesquisa avançada nas bases de dados: foi realizada a busca nas bases da Scopus e WoS do descritivo: “*Machine learning*” AND “*Sentiment analysis*” AND *Touris** e seus respectivos correspondentes em português. Foram aplicados como filtros de pesquisa data de publicação (1984-2023) e apenas artigos. Para a WoS, o descritivo foi pesquisado em todos os campos, retornando 71 artigos. Para a Scopus, o descritivo foi aplicado em título, resumo e palavras-chave, retornando 68 artigos. O período de 1984 a 2023 foi estabelecido como filtro para a pesquisa, uma vez que 1984 marca o início das publicações na área de IA e turismo (Knani et al., 2022).
2. Pesquisa avançada nas bases de dados: foi realizada a busca nas bases da Scopus e WoS do descritivo: “*Machine learning*” AND “*Sentiment analysis*” AND *Touris** e seus respectivos correspondentes em português. Foram aplicados como filtros de pesquisa data de publicação (1984-2023) e apenas artigos. Para a WoS, o descritivo foi pesquisado em todos os campos, retornando 71 artigos. Para a Scopus, o descritivo foi aplicado em título, resumo e palavras-chave, retornando 68 artigos. O período de 1984

a 2023 foi estabelecido como filtro para a pesquisa, uma vez que 1984 marca o início das publicações na área de IA e turismo (Knani et al., 2022).

3. Processamento dos dados: a exportação dos dados foi realizada em 23 de julho de 2024, retornando artigos publicados até essa data. Utilizou-se o Bibliometrix (pacote de linguagem R) para limpeza dos dados e eliminação das duplicatas. Foram obtidos 93 artigos. Feita a leitura dos títulos e resumos para filtragem dos artigos conforme os critérios de exclusão (Quadro 1), obtiveram-se 80 artigos válidos para a pesquisa.
4. Revisão sistemática: foi aplicada a metodologia Methodi Ordinatio para ranqueamento dos 80 artigos (Pagani et al., 2018). Escolheram-se os 10 primeiros artigos do ranking para leitura e discussão.

Quadro 1

Crítérios de inclusão e exclusão

Incluir	Excluir
Artigos focados em turismo, aplicação de ML e AS	Trabalhos fora da área de estudo, trabalhos de eventos, dissertações, teses, livros e outros estudos de revisão
Estudos cuja fontes de dados é o texto de UGC	Estudos que analisam outras formas de UGC, como imagens, áudio e vídeo
Estudos que analisam dados de microblogs e plataformas de avaliação	Estudos que analisam dados de outras mídias sociais, como blogs, Instagram, YouTube etc.
Textos em português, inglês e espanhol	Textos em todas as demais línguas

Nota: Elaboração própria (2026).

O Methodi Ordinatio é uma metodologia que permite classificar os artigos quanto a sua ordem de relevância científica, com base na aplicação da equação InOrdinatio. A equação trabalha com três fatores de um artigo científico: o fator de impacto (IF), o ano da publicação da pesquisa e o número de citações (Ci) (Pagani et al., 2018). Como resultado da fórmula, obtêm-se um indicador de ordem de relevância científica do artigo, permitindo o ranqueamento dos estudos. A InOrdinatio é expressa pela seguinte fórmula:

$$In\ Ordinato = \left(\frac{IF}{1000} \right) + \alpha * [10 - (ResearchYear - PublishYear)] + (\sum Ci)$$

Para a presente pesquisa, o elemento α (que define o grau de importância do ano de publicação na equação) foi definido como 5, justificando-se na relativa incipiência do campo de estudo, mas ainda priorizando estudos recentes. Assim, foram escolhidos para leitura os 10 primeiros artigos do ranking, conforme a Tabela 1. Ressalta-se que, conforme Pagani et al. (2018), o número de estudos selecionados para leitura integral depende exclusivamente de deliberação dos pesquisadores, segundo critérios próprios de acesso aos estudos e disponibilidade de tempo e recursos.

Tabela 1*Ranking dos 10 artigos com base no Methodi Ordinatio*

Ranking	Artigo	Fator impacto	de	Nº Citações	de	Ano publicação	de	InOrdinatio
1	Hussain & Cambria, 2017	5,5		173		2017		218,006
2	Kirilenko et al., 2018	9,7		140		2018		175,009
3	Afzaal et al., 2019	3,9		60		2019		105,004
4	Gour et al., 2020	8,9		44		2020		84,009
5	Kesgin & Murthy, 2019	7,5		39		2019		84,008
6	Chen et al., 2020	2,7		38		2020		83,003
7	Bigne et al., 2021	8,3		40		2021		80,009
8	Kar et al., 2021	0		34		2021		79,000
9	Tao et al., 2019	3,6		38		2019		78,004
10	Li et al., 2022	11,5		52		2022		77,016

Nota: Elaboração própria (2026).

Os resultados são sintetizados e discutidos na seção de Análise dos Resultados dessa pesquisa, a seguir.

ANALISE E DISCUSSAO DOS RESULTADOS

Os principais resultados da análise dos artigos indicam que as publicações se concentram no período de 2017 a 2022. Dois estudos se diferenciam por utilizarem o turismo apenas como contexto de pesquisa, e não como foco principal, como é o caso das investigações voltadas à previsão de sobrevivência de restaurantes e ao desenvolvimento de metodologias para análise de Big Data no campo das ciências sociais (Li et al., 2022; Hussain & Cambria, 2017).

Esses resultados são resumidos no Quadro 2, a seguir.

Quadro 2*Sumário dos resultados da revisão sistemática*

Artigo	Método de ML	Método de treinamento	Método de AS	Outras metodologias	Dados	Avaliação de performance	Foco
Hussain & Cambria, 2017	SVM	Semi-supervisionado	<i>Affective Space</i>	Não	Avaliações: base de dados de pesquisa	Acurácia	Big Data
Kirilenko et al., 2018	SVM, <i>Naïve Bayes</i> e ANN	Supervisionado	ML e léxico (<i>SentiStr enght</i>)	Não	Questionário: percepção de turistas estadunidenses do destino China; avaliações: TripAdvisor; microblog: Twitter	Acurácia, precisão, τ de Kohen, τ de Kendall e recall	Turismo

Artigo	Método de ML	Método de treinamento	Método de AS	Outras metodologias	Dados	Avaliação de performance	Foco
Afzaal et al., 2019	SVM, Naïve Bayes, Random Forest	Múltiplos	ASBA	Árvores de Decisão	Avaliações: sites de viagens	Acurácia, precisão, recall e índice F	Turismo
Gour et al., 2021	SVM e Naïve Bayes	Não supervisionado	Topic Modelling LDA	Clustering	Avaliações: TripAdvisor	Acurácia	Destinos turísticos
Kesgin & Murthy, 2019	Não especificado	Não especificado	Topic Modelling LDA	Regressão linear e regressão logística	Avaliações: Facebook	Correlação	Serviços
Chen et al., 2020	SVM	Supervisionado	TextRank	Microsoft Knowledge Graph	Avaliações: TripAdvisor	Acurácia, precisão, recall e índice F	Turismo
Bigne et al., 2021	Deep Learning e ANN	Não supervisionado	Não especificado	Regressão linear e regressão logística	Avaliações: TripAdvisor	Nenhum	Destinos turísticos
Kar et al., 2021	Não especificado	Não especificado	Topic Modelling LDA	Regressão linear múltipla	Microblog: Twitter	Significância (p-valor)	Serviços
Tao et al., 2019	ANN	Semi-supervisionado	Léxico	Não	Microblog: Sina Weibo	Acurácia, precisão, recall e índice F	Destinos turísticos
Li et al., 2022	Random Forest	Supervisionado	ASBA	Feature importance analysis	Avaliações: sites de restaurantes	C-Index, IBS e feature selection	Serviços

Nota: Elaboração própria (2026).

A pesquisa de Hussain & Cambria (2017) é a mais citada (173 citações). O estudo não trata especificamente do turismo, mas é relevante por sua proposta de comparar a performance de diversos modelos de AS automatizada. O artigo é um dos pioneiros na área, sugerindo o uso de ML semi-supervisionado como melhor método de pesquisa.

Kirilenko et al. (2018) têm o segundo maior número de citações (140) e o segundo maior fator de impacto (9,7). A pesquisa avalia a performance de diferentes classificadores automáticos, comparando-os com classificadores humanos. Conclui que embora os algoritmos supervisionados tenham desempenho comparável ao de classificadores humanos, seus resultados ainda são inferiores em concordância e precisão.

Afzaal et al. (2019) apresentam um framework de Análise de Sentimentos Baseada em Aspectos (ASBA) aplicado a avaliações de turistas de hotéis e restaurantes que usa algoritmos de ML para identificar e classificar aspectos explícitos e implícitos nas avaliações. O framework foi implementado em um app que auxilia turistas a encontrarem os melhores locais com base nos sentimentos classificados.

Gour et al. (2020) propõem uma metodologia multi-método para entender o comportamento dos turistas através de UGC, processando avaliações online sobre destinos

turísticos. Já Kesgin & Murthy (2019) exploram como o conceito de *Social Currency* afeta o engajamento dos consumidores com atrações turísticas através da análise de avaliações online. Seu estudo constata que o conceito tem um impacto positivo nas intenções de revisita e recomendação do destino.

Chen et al. (2020) transformam a AS em um problema de multi-classificação. Os resultados mostram que expansão semântica e técnicas de balanceamento de amostra melhoram a precisão na classificação. Já Bigne et al. (2021) investigam os fatores que influenciam a utilidade percebida das avaliações online de destinos turísticos maduros. O estudo atesta que a experiência do avaliador é um fator consistente em ambas as atrações analisadas na pesquisa.

A pesquisa de Kar et al. (2021) utiliza dados de mídias sociais para identificar fatores que impactam a experiência do consumidor em destinos turísticos na Índia. O estudo propõe um modelo teórico para entender a experiência do consumidor com base em UGC. Tao et al. (2019) investigam como turistas percebem a qualidade do ar em destinos turísticos na China, utilizando dados de redes sociais. O estudo destaca a eficácia da AS para entender as percepções ambientais e propõe implicações para monitoramento e gestão ambiental.

Por fim, o estudo de Li et al. (2022) possui o maior fator de impacto (11,5). A pesquisa utiliza a ASBA para prever a sobrevivência de restaurantes. Os resultados mostram que a ASBA supera a AS tradicional, destacando os sentimentos sobre "serviço" e "sabor" como os mais preditivos dentro do modelo.

O Quadro 3 sumariza as principais metodologias, técnicas e ferramentas utilizadas nos artigos:

Quadro 3

Modelos, técnicas e ferramentas mais utilizados para a pesquisa de AS com ML

Método/Técnica/Ferramenta	Definição	Artigos
Técnicas de PLN (Processamento de Linguagem Natural)	Técnicas que transformam dados textuais em dados que possam ser analisados por algoritmos de forma computacional.	Hussain & Cambria, 2017; Kirilenko et al., 2018; Afzaal et al., 2019; Gour et al., 2020; Kesgin & Murthy, 2019; Chen et al., 2020; Bigne et al., 2021; Kar et al., 2021; Tao et al., 2019; Li et al., 2022
Classificadores baseados em léxicos	Utilizam um dicionário de palavras com sentimentos atribuídos predefinidos.	Kirilenko et al., 2018; Tao et al., 2019
Classificadores baseados em ML	Aprende a classificar os sentimentos com base nos próprios dados.	Hussain & Cambria, 2017; Kirilenko et al., 2018; Afzaal et al., 2019; Gour et al., 2020; Kesgin & Murthy, 2019; Chen et al., 2020; Bigne et al., 2021; Kar et al., 2021; Tao et al., 2019; Li et al., 2022
<i>Topic Modelling</i>	Técnica avançada de PLN que busca identificar os tópicos que resultam no sentimento classificado.	Gour et al., 2020; Kesgin & Murthy, 2019; Kar et al., 2021
ASBA	Classifica não apenas o texto, mas diversos aspectos presentes no texto, extraindo mais de um tema tratado no mesmo texto.	Afzaal et al., 2019; Li et al., 2022

Método/Técnica/Ferramenta	Definição	Artigos
<i>Clusteing</i>	Técnica de agrupamento dos textos.	Gour et al., 2020
Regressão linear múltipla	Técnica estatística usada para modelar a relação entre uma variável dependente e várias variáveis independentes. Busca prever o valor da variável dependente com base nas variáveis independentes.	Kesgin & Murthy, 2019; Bigne et al., 2021; Kar et al., 2021
Regressão logística	Técnica estatística usada para modelar a probabilidade de um evento ocorrer, podendo ser binário (sim ou não).	Kesgin & Murthy, 2019; Bigne et al., 2021
Algoritmos de <i>Naïve Bayes</i>	Abordagem probabilística que estima a probabilidade de certo sentimento estar presente no texto.	Kirilenko et al., 2018; Afzaal et al., 2019; Gour et al., 2020
Algoritmos de <i>Support Vector Machine</i> (SVM)	Abordagem matemática que procura a melhor separação linear entre os dados que expressam sentimentos positivos e negativos.	Hussain & Cambria, 2017; Kirilenko et al., 2018; Afzaal et al., 2019; Gour et al., 2020; Chen et al., 2020
Algoritmos <i>Artificial Neural Networks</i> (ANN) ou Redes Neurais Artificiais	Abordagem que mimetiza um cérebro biológico ao processar os dados através de uma rede de neurônios auto-organizada.	Kirilenko et al., 2018; Afzaal et al., 2019; Bigne et al., 2021; Tao et al., 2019
Árvores de Decisão	É composta por nós que fazemos perguntas ou condições sobre os dados. Cada ramo da árvore corresponde a uma resposta (como "sim" ou "não"), e as ramificações continuam até que se atinja uma folha, que contém a classificação final.	Afzaal et al., 2019; Li et al., 2022
<i>Random Forest</i>	Múltiplas árvores de decisão a partir de diferentes amostras de dados. Cada árvore faz uma previsão de sentimento e, no final, o modelo toma uma decisão baseada no voto da maioria das árvores.	Afzaal et al., 2019; Li et al., 2022

Nota: Elaboração própria (2026).

Os estudos demonstram que embora o ML seja amplamente utilizado na automação da AS, os modelos requerem técnicas complementares para cumprir a tarefa de classificação.

Enquanto os algoritmos de ML classificam a polaridade do sentimento, ferramentas de PLN extraem e organizam os dados, convertendo textos em formatos computacionais adequados para o processamento textual (Hussain & Cambria, 2017; Bigne et al., 2021; Kirilenko et al., 2018; Chen et al., 2020; Kar et al., 2021). Dentre as ferramentas de PLN, destacam-se o *Topic Modelling* e a ASBA.

O *Topic Modelling* é uma técnica computacional que identifica o principal tópico (como preço, segurança ou lazer) associado a determinando sentimento (Gour et al., 2020; Kesgin & Murthy, 2019; Kar et al., 2021). Já a ASBA detecta múltiplos tópicos (chamados aspectos) associados ao mesmo sentimento, permitindo análise detalhada e mais precisa dos dados (Afzaal et al., 2019; Li et al., 2022).

Somente após passar por esses processos o texto pode ser submetido ao algoritmo de ML. Dos modelos utilizados, sobressaem-se os algoritmos de SVM e *Naïve Bayes*. Ambos são os mais recomendados pelas pesquisas que comparam diferentes metodologias (Hussain & Cambria, 2017; Kirilenko et al., 2018; Afzaal et al., 2019).

SVM (*Support Vector Machine*) é uma abordagem matemática que procura a melhor separação linear entre os dados que expressam sentimentos positivos e negativos. Dos modelos que empregam SVM, o algoritmo semi-supervisionado de Hussain & Cambria (2017) obteve 63% de acurácia na classificação de sentimentos. O modelo proposto pelos pesquisadores alcançou 88,5% de acurácia quando aplicado a uma base de dados do site Rotten Tomatoes. Comparativamente, na pesquisa de Kirilenko et al. (2018) para classificação de avaliações do TripAdvisor, seu algoritmo supervisionado obteve 87% e 88% de acurácia e 52% e 42% de precisão em comparação com dois classificadores humanos. Já o modelo de Afzaal et al. (2019) obteve 87,02% de acurácia e 87% de precisão ao classificar avaliações de restaurantes, e 89,93% de acurácia e 89% de precisão para classificação de avaliações de hotéis.

O algoritmo de *Naïve Bayes* estima a probabilidade de certo sentimento estar presente num texto (Hussain & Cambria, 2017; Kirilenko et al. 2018; Afzaal et al. 2019; Gour et al., 2020). Ao ser testado junto a outros modelos na classificação de respostas de um questionário eletrônico, o algoritmo de Kirilenko et al. (2018) obteve 62% e 74% de acurácia em comparação com dois classificadores humanos. No estudo de Afzaal et al. (2019), o algoritmo obteve 88,08% de acurácia e 88% de precisão para classificação de avaliações de restaurantes, e 90,53% de acurácia e 90% de precisão para avaliações de hotéis, sendo o melhor desempenho dentre os modelos testados.

Tanto SVM quanto *Naïve Bayes* são considerados algoritmos simples. Técnicas mais avançadas e complexas ainda não entregam a mesma performance desses algoritmos. Entre essas técnicas, destacam-se o uso de ANN, modelo que tem por objetivo mimetizar um cérebro biológico ao processar os dados através de uma rede de neurônios auto-organizada (Kirilenko et al., 2018).

Quanto às formas de treinamento, há uma preferência por algoritmos de ML treinados de forma supervisionada ou semi-supervisionada (Hussain & Cambria, 2017; Kirilenko et al., 2018; Chen et al., 2020; Tao et al., 2019; Li et al., 2022), com apenas dois estudos (Bigne et al., 2021; Gour et al, 2020) optando pelo treinamento não-supervisionado.

Discussão dos Dados

Conforme a leitura dos artigos, predominam dois tipos de estudo: aquele cujo foco está em comparar e/ou desenvolver algoritmos para pesquisa (Hussain & Cambria, 2017; Kirilenko et al., 2018; Afzaal et al., 2019; Gour et al., 2020; Chen et al., 2020); e aquele que aplica o ML para AS como uma de suas ferramentas de pesquisa, visando o respectivo objetivo proposto pelo estudo (Kesgin & Murthy, 2019; Bigne et al., 2021; Kar et al, 2021; Tao et al., 2019; Li et al., 2022). Os artigos do primeiro tipo são os mais citados, visto que suas pesquisas muitas vezes servem de base metodológica para pesquisas mais aplicadas, do segundo tipo. Essa divisão em dois tipos permite responder ao primeiro objetivo específico da pesquisa: compreender a aplicação do Aprendizado de Máquina nos estudos em turismo.

Exemplos do primeiro tipo incluem: Kirilenko et al. (2018) comparam métodos de AS no turismo com a performance humana, concluindo que no ML os algoritmos de SVM e *Naïve Bayes* são os que demonstram melhor performance; Afzaal et al. (2019) desenvolvem um framework de pesquisa com aplicação de ML para ASBA, comparando mais de um modelo de ML e propondo ainda a aplicação do framework em um app móvel; e Gour et al. (2020) propõe uma framework multi-método com uso de ML não-supervisionado e *Topic Modelling LDA* para pesquisas com UGC.

No segundo tipo destacam-se: Kesgin & Murthy (2019) aplicam ML, *Topic Modelling LDA* e regressão linear em UGC para identificar a relação entre lealdade do turista, avaliações online e *Social Currency*; Kar et al. (2021) aplicam ML e *Topic Modelling LDA* em conjunto com regressão linear para identificar em posts de microblogs os fatores que influenciam a experiência do turista no destino; e Tao et al. (2019) combinam o uso de léxicos com ANN para avaliar em posts de microblogs a percepção da qualidade do ar em destinos turísticos.

Embora com aplicações práticas evidentes, os estudos sobre ML e AS também abordam questões teóricas, destacando-se a Computação Afetiva como eixo comum, mesmo quando não mencionada explicitamente (Hussain & Cambria, 2017; Kirilenko et al., 2018; Afzaal et al., 2019; Chen et al., 2020; Gour et al., 2021; Kar et al., 2021; Li et al., 2022). Essa área interdisciplinar investiga como máquinas interpretam e simulam emoções humanas, aproximando a interação homem-máquina (Kumar, 2021; Wang et al., 2022; Afzal et al., 2024), reunindo psicologia, ciência da computação, robótica e sociologia (Arya et al., 2021; Devillers & Cowie, 2023).

Avanços em ML e AS são fundamentais para robôs, chatbots e IA generativa, visando interações mais empáticas e humanizadas (Arya et al., 2021; Kumar, 2021; Afzal et al., 2024; Devillers & Cowie, 2023; Wang et al., 2022). No turismo, a Computação Afetiva ainda é pouco explorada, mas tende a ganhar espaço. Estudos recentes buscam alinhar inovação tecnológica à criação de experiências significativas, com destaque para o papel da IoT, das mídias sociais e da própria cocriação de experiências (Xiang et al., 2021; Herath & Mittal, 2022; Gursoy et al., 2024).

Nesse contexto, a Computação Afetiva pode transformar a tecnologia em um recurso de hospitalidade, promovendo interações mais personalizadas e emocionalmente ricas entre turistas e destinos. A identificação dessas lacunas responde ao segundo objetivo específico da pesquisa: identificar possíveis lacunas nas pesquisas do uso de Aprendizado de Máquina no turismo.

Ao analisar os resultados da revisão sistemática foi possível responder as questões de pesquisa.

P1: quais as vantagens e desvantagens dos principais modelos de ML utilizados para Análise de Sentimentos no turismo?

SVM e *Naïve Bayes* se destacam por sua simplicidade em comparação a outros algoritmos. O SVM tem como vantagens a alta precisão em classificação, a efetividade em dados de texto vetoriais e o bom desempenho mesmo com amostras pequenas de dados. Suas desvantagens incluem o alto custo computacional quando se lida com grandes volumes de dados e a difícil interpretação dos resultados. Já o *Naïve Bayes* é simples e veloz, tornando-o ideal para grandes volumes de dados, além de apresentar bom desempenho em textos curtos e

análises simples (como a polaridade positivo/negativo). Entretanto, ele possui menor precisão em análises complexas com nuances emocionais.

P2: quais as metodologias empregadas nas principais pesquisas no tema?

A maioria dos estudos aplica a aprendizagem supervisionada ou semi-supervisionada. Dentre os modelos que seguem esse método de aprendizagem, destacam-se algoritmos de SVM, *Naïve Bayes* e ANN para classificar sentimentos. Dentre as técnicas para extração do dado textual, destacam-se o uso de *Topic Modelling LDA* e de ABSA. Os estudos também combinam a AS com outras análises estatísticas, combinando dados de sentimento com variáveis geográficas e temporais para identificar padrões de percepção, e identificando relações entre os sentimentos e outras variáveis através de análises fatoriais.

P3: quais os principais campos de pesquisa inseridos nesse tema?

A gestão de destinos turísticos é um dos principais campos de pesquisa inseridos no tema. O planejamento de destinos turísticos se beneficia do uso de AS para identificar percepções de turistas e ajustar estratégias de marketing e desenvolvimento. A Computação Afetiva é outro tema que aparece com recorrência, mesmo de que forma indireta. Nela, a integração da análise de emoções com UGC é usada para compreender melhor a percepção da experiência turística nos destinos. As pesquisas também se voltam para campos como marketing turístico, satisfação do turista, qualidade da experiência turística, qualidade de serviços, comportamento do consumidor, além do estudo de metodologias de pesquisa e desenvolvimento de algoritmos e frameworks.

P4: quais as tendências desse tema e como elas se relacionam com outros campos e temas de pesquisa?

As tendências em ML e AS no turismo apontam para uma automação mais avançada, com destaque para técnicas de *Deep Learning* (DL), modelos mais complexos que permitem análises mais precisas e contextuais. Há também um crescimento no uso de IA para análises preditivas e estatísticas, possibilitando a antecipação de comportamentos e preferências dos turistas. A ampliação das fontes de dados, incluindo diversas mídias sociais e dados não estruturados, reforça essa evolução. No campo da sustentabilidade, tende a crescer o interesse em avaliar percepções sobre práticas sustentáveis e impactos ambientais, enquanto frameworks personalizados vêm sendo desenvolvidos para contextos regionais específicos.

Esses avanços impactam múltiplas áreas do turismo. Na sustentabilidade, a AS ajuda a entender melhor a visão dos turistas sobre questões ambientais, contribuindo para políticas mais eficazes. No marketing, permite identificar preferências e segmentar campanhas com maior precisão. Já no planejamento e na gestão pública, a análise baseada em UGC fornece suporte estratégico para decisões mais informadas. Além disso, a modelagem da experiência e as análises preditivas auxiliam na previsão de demanda e na otimização da oferta de serviços turísticos.

A aplicação dessas metodologias ultrapassa o campo do turismo e se relaciona com outras áreas, como a psicologia e a sociologia, que utilizam a AS para estudar o comportamento humano no ambiente digital. Ademais, a capacidade de interpretar emoções e sentimentos de forma computacional tem um impacto na interação entre humanos e máquinas, promovendo avanços em tecnologias que buscam tornar a experiência online mais intuitiva e humanizada.

Destarte, a partir dessa discussão é possível elaborar a agenda de pesquisa futura para o campo de estudo e responder ao objetivo de pesquisa.

1. Investir em desenvolvimento e aplicação de algoritmos mais avançados: tecnologias de DL, ANN e *Random Forest* se mostram mais flexíveis e capazes de lidar com problemas de maior complexidade. Entretanto, ainda apresentam empecilhos para sua aplicação prática na pesquisa, como a exigência de grandes volumes de dados e alta capacidade de processamento computacional, além de apresentarem resultados muitas vezes difíceis para interpretação. Tais obstáculos seriam superados com investimento maior nessas tecnologias e em sua aplicação para a pesquisa científica em turismo. De um lado, o avanço da tecnologia costuma simplificar seu desenvolvimento, do outro, seu ensino e disseminação a tornam mais acessível.
2. Ampliar e aprofundar a mensuração de emoções: as pesquisas, em sua maioria, restringem a AS a uma escala que vai do positivo ao negativo, somente. Mas sentimentos podem ser analisados de maneira mais complexa, ampliando as discussões possíveis numa pesquisa. A análise emocional traz não apenas uma escala maior, mas também um espectro de emoções, incluindo tristeza, alegria, raiva, apatia, dentre outras, cada qual com suas nuances. A adoção desse espectro nas pesquisas possibilitaria maior entendimento do papel das emoções e sentimentos no turismo. Nesse contexto, a pesquisa poderia ser ampliada para análises multimodais, integrando a análise sentimental de textos, imagens e vídeos, desenvolvendo assim modelos altamente precisos nos mais variados contextos.
3. Aplicar outros métodos estatísticos em conjunto com a AS: o uso de outros métodos estatísticos expande a contribuição de uma pesquisa, adicionando mais complexidade aos seus resultados. Técnicas de modelagem de dados, estatística preditiva, *clustering* e outros métodos podem ser combinados com os resultados de uma AS para trazer mais do que apenas os sentimentos de turistas expressos online, mas apontar para investigações mais profundas de perfil, percepção, satisfação etc.
4. Aprofundar a relação com outros conceitos, campos e áreas de pesquisa: a AS produz resultados que sozinhos oferecem poucas contribuições ao campo. Sua relevância cresce quando aplicada a contextos específicos ou em diálogo com outros temas e campos de pesquisa. Marketing, planejamento turístico e imagem de destino são áreas que se beneficiam da maior compreensão emocional da demanda turística, por exemplo. Pesquisas podem investigar como determinadas variáveis afetam a demanda turística, analisar a recepção de eventos e festividades, e ajudar a investigar problemas atuais do turismo, como a turismofobia e o *overtourism*.
5. Aprofundar a relação entre turismo e Computação Afetiva: o turismo é sempre transformado pela tecnologia, avançando lado a lado com o progresso tecnológico. Um entendimento e alinhamento maior com a Computação Afetiva permite ampliar a inserção da IA e suas tecnologias de forma benéfica para o turismo. Ao investir em tecnologias que ampliem a interpretação de emoções pelas máquinas, as soluções inteligentes e tecnológicas presentes no turismo se tornam cada vez mais acolhedoras e receptivas. Uma prioridade seria desenvolver modelos teóricos que integrem Turismo, Marketing, IA e Psicologia.
6. Disponibilizar aos turistas e ao residente os resultados de pesquisas: a acessibilidade é um fator importante para a ciência. Os resultados de uma AS muitas vezes ficam restritos a estudos de caso e às instituições de pesquisa. Dever-se-ia explorar formas de transformar esses resultados em informação para o turista e o residente do destino. Por

exemplo, uma solução que permitisse ao turista verificar quais as emoções e sentimentos que determinada experiência turística provoca.

Com essa proposta, espera-se direcionar as futuras pesquisas no campo de Aprendizado de Máquina aplicado à Análise de Sentimentos para o turismo. A seguir, segue-se para as conclusões da pesquisa.

CONCLUSÃO

O presente estudo respondeu sua principal pergunta de pesquisa. Foi identificado que os principais modelos de ML utilizados no Turismo para AS a partir do UGC são algoritmos supervisionados ou semi-supervisionados, com aplicação baseada em *Naïve Bayes* ou SVM e utilização de *Topic Modelling LDA* ou ASBA para mineração dos sentimentos no texto. Assim, após análise e discussão dos resultados a pesquisa cumpriu com seu principal objetivo, isto é, elaborar uma agenda de pesquisa futura de ML para a pesquisa com AS com UGC em turismo. O estudo alcançou seus objetivos específicos e respondeu às perguntas de pesquisa. Foi possível compreender a utilização do ML nos estudos em turismo e identificar possíveis lacunas nas pesquisas do uso de ML no turismo.

O Aprendizado de Máquina tem potencial para lidar com volumes massivos de dados no setor turístico, e, aliado às técnicas de Análise de Sentimentos, amplia significativamente as possibilidades de exploração de dados secundários de UGC. Considerando que o turismo é uma atividade dependente da informação, essas metodologias oferecem contribuições relevantes, gerando novos conhecimentos que podem qualificar tanto a gestão quanto a experiência turística. Pesquisas futuras devem concentrar esforços no desenvolvimento de modelos mais complexos e contextualmente adaptados, além de aprofundar a relação teórica entre o turismo e a Computação Afetiva, que se mostra promissora para o aprimoramento das interações entre turistas e tecnologias. Ademais, é fundamental que os avanços obtidos sejam disseminados, tanto para os profissionais do setor quanto para os turistas, de modo a potencializar sua aplicação prática e promover um turismo mais inteligente, responsivo e centrado na experiência humana.

O estudo foi limitado a publicações em inglês, espanhol e português. Futuras pesquisas podem investigar publicações em outras línguas, especialmente em mandarim e as publicações chinesas. Esta revisão sistemática incluiu apenas os 10 primeiros artigos obtidos com o Methodi Ordinatio. Sugere-se leitura dos demais artigos do ranking para futuras pesquisas, ou o uso de outras metodologias de revisão sistemática.

REFERENCIAS

Cândido, G. A., & Zucco, F. M. (2021). Turismo de base comunitária e desenvolvimento local: Estudo de caso em comunidades rurais do Rio Grande do Sul. *Revista Brasileira de Pesquisa em Turismo*, 15(1), 45–62. <https://doi.org/10.7784/rbtur.v15i1.2004>

- Afzaal, M., Usman, M., & Fong, A. (2019). Tourism Mobile App with Aspect-Based Sentiment Classification Framework for tourist reviews. *IEEE Transactions on Consumer Electronics*, 65(2), 233–242. <https://doi.org/10.1109/tce.2019.2908944>
- Afzal, S., Khan, H. A., Piran, M. J., & Lee, J. W. (2024). A comprehensive survey on affective computing: challenges, trends, applications, and future directions. *IEEE Access*, 12, 96150–96168. <https://doi.org/10.1109/access.2024.3422480>
- Alaei, A. R., Becken, S., & Stantic, B. (2017). Sentiment Analysis in Tourism: Capitalizing on big data. *Journal of Travel Research*, 58(2), 175–191. <https://doi.org/10.1177/0047287517747753>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Arya, R., Singh, J., & Kumar, A. (2021). A survey of multidisciplinary domains contributing to affective computing. *Computer Science Review*, 40, 100399. <https://doi.org/10.1016/j.cosrev.2021.100399>
- Barbera, G., Araújo, L., & Fernandes, S. (2023). The Value of web Data Scraping: An application to TripAdvisor. *Big Data and Cognitive Computing*, 7(3), 121. <https://doi.org/10.3390/bdcc7030121>
- Bigne, E., Ruiz, C., Cuenca, A., Perez, C., & Garcia, A. (2021). What drives the helpfulness of online reviews? A deep learning study of sentiment analysis, pictorial content and reviewer expertise for mature destinations. *Journal of Destination Marketing & Management*, 20, 100570. <https://doi.org/10.1016/j.jdmm.2021.100570>
- Calderón-Fajardo, V., Anaya-Sánchez, R., & Molinillo, S. (2024). Understanding destination brand experience through data mining and machine learning. *Journal of Destination Marketing & Management*, 31, 100862. <https://doi.org/10.1016/j.jdmm.2024.100862>
- Chen, W., Xu, Z., Zheng, X., Yu, Q., & Luo, Y. (2020). Research on Sentiment Classification of Online Travel Review text. *Applied Sciences*, 10(15), 5275. <https://doi.org/10.3390/app10155275>
- Chen, J., Becken, S., & Stantic, B. (2022). Assessing destination satisfaction by social media: An innovative approach using Importance-Performance Analysis. *Annals of Tourism Research*, 93, 103371. <https://doi.org/10.1016/j.annals.2022.103371>
- Cruz, M. M. D. (2018). *Curitiba: um destino turístico inteligente?* [Dissertação]. Universidade Federal do Paraná. <https://acervodigital.ufpr.br/handle/1884/55445>
- Devillers, L., & Cowie, R. (2023). Ethical Considerations on Affective Computing: An Overview. *Proceedings of the IEEE*, 111(10), 1445–1458. <https://doi.org/10.1109/jproc.2023.3315217>
- European Commission: Joint Research Centre, Delipetrev, B., Tsinaraki, C. and Kostić, U., *AI watch, historical evolution of artificial intelligence – Analysis of the three main paradigm shifts in AI*, Publications Office, 2020, <https://data.europa.eu/doi/10.2760/801580>
- Gonçalves, G. da Costa, & Oliveira, F. Tavres de (2023). O uso do Data Science na análise dos indicadores do Turismo: uma revisão das pesquisas no Brasil (2012-2022). *Revista Iberoamericana De Turismo*, 13(1). <https://www.each.usp.br/turismo/publicacoesdeturismo/ref.php?id=16542>
- Gour, A., Aggarwal, S., & Erdem, M. (2021). Reading between the lines: analyzing online reviews by using a multi-method Web-analytics approach. *International Journal of Contemporary Hospitality Management*, 33(2), 490–512. <https://doi.org/10.1108/ijchm-07-2020-0760>

- Gursoy, D., Luongo, S., Della Corte, V., & Sepe, F. (2024). Smart tourism destinations: an overview of current research trends and a future research agenda. *Journal of Hospitality and Tourism Technology*. <https://doi.org/10.1108/jhtt-10-2023-0339>
- Herath, H., & Mittal, M. (2022). Adoption of artificial intelligence in smart cities: A comprehensive review. *International Journal of Information Management Data Insights*, 2(1), 100076. <https://doi.org/10.1016/j.ijime.2022.100076>
- Hussain, A., & Cambria, E. (2017). Semi-supervised learning for big social data analysis. *Neurocomputing*, 275, 1662–1673. <https://doi.org/10.1016/j.neucom.2017.10.010>
- Hu, M., Li, M., Chen, Y., & Liu, H. (2025). Tourism forecasting by mixed-frequency machine learning. *Tourism Management*, 106, 105004. <https://doi.org/10.1016/j.tourman.2024.105004>
- Kar, A. K., Kumar, S., & Ilavarasan, P. V. (2021). Modelling the service experience encounters using User-Generated Content: a text mining approach. *Global Journal of Flexible Systems Management*, 22(4), 267–288. <https://doi.org/10.1007/s40171-021-00279-5>
- Kesgin, M., & Murthy, R. S. (2019). Consumer engagement: the role of social currency in online reviews. *Service Industries Journal*, 39(7–8), 609–636. <https://doi.org/10.1080/02642069.2018.1553237>
- Kirilenko, A. P., Stepchenkova, S. O., Kim, H., & Li, X. (2018). Automated Sentiment Analysis in Tourism: Comparison of approaches. *Journal of Travel Research*, 57(8), 1012–1025. <https://doi.org/10.1177/0047287517729757>
- Kirtil, İ. G., & Aşkun, V. (2021). Artificial Intelligence in Tourism: A Review and Bibliometrics research. *Advances in Hospitality and Tourism Research (AHTR)*, 9(1), 205–233. <https://doi.org/10.30519/ahtr.801690>
- Knani, M., Echchakoui, S., & Ladhari, R. (2022). Artificial intelligence in tourism and hospitality: Bibliometric analysis and research agenda. *International Journal of Hospitality Management*, 107, 103317. <https://doi.org/10.1016/j.ijhm.2022.103317>
- Kumar, S. (2021). Deep learning based affective computing. *Journal of Enterprise Information Management*, 34(5), 1551–1575. <https://doi.org/10.1108/jeim-12-2020-0536>
- Leelawat, N., Jariyapongpaiboon, S., Promjun, A., Boonyarak, S., Saengtabtim, K., Laosunthara, A., Yudha, A. K., & Tang, J. (2022). Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning. *Heliyon*, 8(10), e10894. <https://doi.org/10.1016/j.heliyon.2022.e10894>
- Li, H., Yu, B. X., Li, G., & Gao, H. (2022). Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews. *Tourism Management*, 96, 104707. <https://doi.org/10.1016/j.tourman.2022.104707>
- Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, 68, 301–323. <https://doi.org/10.1016/j.tourman.2018.03.009>
- Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). A review of Text Corpus-Based Tourism Big Data Mining. *Applied Sciences*, 9(16), 3300. <https://doi.org/10.3390/app9163300>
- Lima, J. I. M. P. de, Pessanha, G. R. G., De Araújo, M. V. P., De Almeida Alves, R. C., Cesário, M. F. P., & De Oliveira Rocha, A. C. F. (2022). Marca lugar Pernambuco: análise dos sentimentos dos usuários através das hashtags do Instagram. *ReMark - Revista Brasileira De Marketing*, 21(1), 154–184. <https://doi.org/10.5585/remark.v21i1.20578>
- Lin, B., Zhang, Y. C., & Lee, W. (2024). Tourists' perceptions of proximity tourism: Insights from sentiment analysis and fsQCA. *Journal of Outdoor Recreation and Tourism*, 49, 100848. <https://doi.org/10.1016/j.jort.2024.100848>

- Liu, J., Hu, S., Mehraliyev, F., Zhou, H., Yu, Y., & Yang, L. (2023). Recognizing emotions in restaurant online reviews: a hybrid model integrating deep learning and a sentiment lexicon. *International Journal of Contemporary Hospitality Management*, 36(9), 2955–2976. <https://doi.org/10.1108/ijchm-02-2023-0244>
- Lyu, J., Khan, A., Bibi, S., Chan, J. H., & Qi, X. (2022). Big data in action: An overview of big data studies in tourism and hospitality literature. *Journal of Hospitality and Tourism Management*, 51, 346–360. <https://doi.org/10.1016/j.jhtm.2022.03.014>
- Maia, C. R., Maçada, A. C. G., Lunardi, G. L., & Sengik, A. R. (2024). Data monetization capability: from the systematic literature review to the proposition of a multidimensional model. *Revista de Administração da UFSM*, 17(2), e6. <https://doi.org/10.5902/1983465986416>
- Manosso, F. C., & Ruiz, T. C. D. (2021). Using sentiment analysis in tourism research: A systematic, bibliometric, and integrative review. *HAL (Le Centre Pour La Communication Scientifique Directe)*. <https://doi.org/10.5281/zenodo.5548426>
- Marigliano, P. (2024). Analyzing tourism reviews using Deep Learning and AI to predict sentiments. *Quality & Quantity*. <https://doi.org/10.1007/s11135-024-01840-x>
- Mountasser, I., Ouhbi, B., Frikh, B., & Hdioud, F. (2020). Big data research in the tourism industry. *International Journal of Mobile Computing and Multimedia Communications*, 11(4), 26–41. <https://doi.org/10.4018/ijmcmc.2020100102>
- Neratika, R. R. (2016). *O uso de indicadores em turismo: uma ferramenta para analisar a evolução dos destinos turísticos inteligentes* [Trabalho de Graduação]. Universidade Federal do Paraná.
- Nieves-Pavón, S., López-Mosquera, N., & Jiménez-Naranjo, H. (2024). The role emotions play in loyalty and WOM intention in a Smart Tourism Destination Management. *Cities*, 145, 104681. <https://doi.org/10.1016/j.cities.2023.104681>
- Nnachi, A. B., Arinze, E. D., & Uchekchukwu, A. J. (2024). Exploring the Frontiers of Data Analysis: A Comprehensive review. *INOSR APPLIED SCIENCES*, 12(1), 62–68. <https://doi.org/10.59298/inosras/2024/12.1.62680>
- Pagani, R. N., Kovalski, J. L., & Resende, L. M. (2015). Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication. *Scientometrics*, 105(3), 2109–2135. <https://doi.org/10.1007/s11192-015-1744-x>
- Porto, A. P. D. S., Júnior, V. N. F., & Pozzebon, E. (2024). Contribuições Da Inteligência Artificial Para o Turismo. *Revista Rosa Dos Ventos - Turismo E Hospitalidade*, 15(4), 1028–1042. <https://doi.org/10.18226/21789061.v15i4p1028>
- Romele, A., Severo, M., & Furia, P. (2018). Digital hermeneutics: from interpreting with machines to interpretational machines. *AI & Society*, 35(1), 73–86. <https://doi.org/10.1007/s00146-018-0856-2>
- Santos, V. S. D., De Sousa, S. J. A., Santos, L. M. L., Filho, L. a. M. M., De Santana Porte, M., Da Silva Taveira, M., & De Oliveira Alexandre, M. L. (2024). Inteligência Artificial nos estudos e pesquisas em Turismo no Brasil. *Revista Brasileira De Pesquisa Em Turismo*, 18, 2896. <https://doi.org/10.7784/rbtur.v18.2896>
- Sari, B. A., Alkhalidi, R., Alsaffar, D., Alkhalidi, T., Almaymuni, H., Alnaim, N., Alghamdi, N., & Olatunji, S. O. (2022). Sentiment analysis for cruises in Saudi Arabia on social media platforms using machine learning algorithms. *Journal of Big Data*, 9(1). <https://doi.org/10.1186/s40537-022-00568-5>
- Schank, R. C. (1991). Where's the AI? *AI Magazine*, 12(4), 38–49. <https://doi.org/10.1609/aimag.v12i4.917>
- Silva, N. C. da. (2023). *Big Data e demanda turística em destinos turísticos: uma análise do perfil dos visitantes da Região Sul do Brasil* [Dissertação de Mestrado]. Universidade

- Federal do Paraná. Disponível em:
<https://acervodigital.ufpr.br/xmlui/bitstream/handle/1884/83881/R%20-%20D%20-%20NATALIA%20CRISTINA%20DA%20SILVA.pdf?sequence=1&isAllowed=y>
- Song, H., Qiu, R. T., & Park, J. (2023). Progress in tourism demand research: Theory and empirics. *Tourism Management*, 94, 104655.
<https://doi.org/10.1016/j.tourman.2022.104655>
- Tao, Y., Zhang, F., Shi, C., & Chen, Y. (2019). Social Media Data-Based Sentiment Analysis of Tourists' air quality Perceptions. *Sustainability*, 11(18), 5070.
<https://doi.org/10.3390/su11185070>
- Troisi, O., Visvizi, A., & Grimaldi, M. (2023). Digitalizing business models in hospitality ecosystems: toward data-driven innovation. *European Journal of Innovation Management*, 26(7), 242–277. <https://doi.org/10.1108/ejim-09-2022-0540>
- Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. *Journal of Big Data*, 2(1). <https://doi.org/10.1186/s40537-015-0030-3>
- Tsaih, R., & Hsu, C. C. (2018). Artificial Intelligence in Smart Tourism: A Conceptual Framework. *AIS Electronic Library*, 124–133. <https://aisel.aisnet.org/iceb2018/89/>
- Veal, A. J. (2011). *Metodologia de pesquisa em lazer e turismo*. Aleph.
- Wang, X., Zhu, F., Tang, J., Shen, L., & Liu, D. (2021). Applications, Experiences, and Challenges of Smart Tourism Development in China. *the Journal of Urban Technology/Journal of Urban Technology*, 29(4), 101–126.
<https://doi.org/10.1080/10630732.2021.1879605>
- Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., Gao, S., Sun, Y., Ge, W., Zhang, W., & Zhang, W. (2022). A systematic review on affective computing: emotion models, databases, and recent advances. *Information Fusion*, 83–84, 19–52.
<https://doi.org/10.1016/j.inffus.2022.03.009>
- Weismayer, C., Gunter, U., & Önder, I. (2021). Temporal variability of emotions in social media posts. *Technological Forecasting and Social Change*, 167, 120699.
<https://doi.org/10.1016/j.techfore.2021.120699>
- Wu, J., You, H., & Du, J. (2025). AI generations: from AI 1.0 to AI 4.0. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1585629>
- Xiang, Z., Stienmetz, J., & Fesenmaier, D. R. (2021). Smart Tourism Design: Launching the annals of tourism research curated collection on designing tourism places. *Annals of Tourism Research*, 86, 103154. <https://doi.org/10.1016/j.annals.2021.103154>
- Xu, Q. A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, 3, 100073. <https://doi.org/10.1016/j.dajour.2022.100073>
- Zang, Z., Fu, H., Cheng, J., Raza, H., & Fang, D. (2024). Digital threads of architectural heritage: navigating tourism destination image through social media reviews and machine learning insights. *Journal of Asian Architecture and Building Engineering*, 1–18. <https://doi.org/10.1080/13467581.2024.2399681>

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio, por meio de bolsas e/ou recursos financeiros, essenciais para a realização desta pesquisa.

DECLARAÇÃO DE DISPONIBILIDADE DOS DADOS

Todos os dados relevantes estão disponíveis no texto.

INFORMAÇÕES DO(S) AUTOR(ES)

1

Mestrando em Turismo pela Universidade Federal do Paraná (UFPR). Bacharel em Turismo pela Universidade Federal do Paraná (UFPR). E-mail: bruno.homann@ufpr.br | ORCID: <https://orcid.org/0000-0001-6121-6479>

Contribuições: Conceituação, Curadoria de dados, Análise formal, Investigação, Metodologia, Administração do projeto, Supervisão, Validação, Redação do manuscrito original, Redação - revisão e edição.

2

Professora e pesquisadora permanente do Programa de Pós-Graduação em Turismo da Universidade Federal do Paraná (UFPR). Doutora em Turismo e Hotelaria pela Universidade do Vale do Itajaí (UNIVALI). Mestra em Turismo e Hotelaria pela Universidade do Vale do Itajaí (UNIVALI). Bacharel em Turismo pela Universidade Federal de Pelotas (UFPEL). E-mail: melisepereira@ufpr.br | ORCID: <https://orcid.org/0000-0003-1904-8429>

Contribuições: Conceituação, Curadoria de dados, Análise formal, Investigação, Metodologia, Administração do projeto, Supervisão, Validação, Redação do manuscrito original, Redação - revisão e edição.

REVISTA VINCULADA A



**UNIVERSIDADE FEDERAL
DE MATO GROSSO DO SUL**