

AVALIAÇÃO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL PARA A PREDIÇÃO DE EXPERIÊNCIAS TURÍSTICAS

EVALUATION OF ARTIFICIAL INTELLIGENCE METHODS FOR PREDICTING TOURIST EXPERIENCES

Recebido em 01.10.2025 Aprovado em 24.03.2026

Avaliado pelo sistema double blind review

DOI: <https://doi.org/10.12712/rpca.v.194.69407>

Denilson Lopes

denilson@ufpr.br

Universidade Federal do Paraná - Curitiba/Paraná, Brasil

<https://orcid.org/0009-0002-6811-7275>

João Eugenio Marynowski

jeugenio@ufpr.br

Universidade Federal do Paraná - Curitiba/Paraná, Brasil

<https://orcid.org/0000-0002-0168-7217>

José Elmar Feger

elmar@ufpr.br

Universidade Federal do Paraná - Curitiba/Paraná, Brasil

<https://orcid.org/0000-0002-1982-4179>

Resumo

Este artigo avalia a aplicação de Inteligência Artificial na análise de comentários de turistas para melhorar os serviços turísticos. Foram testados métodos de Aprendizado de Máquina (Naïve Bayes, Redes Neurais, SVM e LSTM) para a classificação e predição de experiências a partir de comentários da plataforma TripAdvisor, envolvendo os parques Jalapão e Superagui. O Naïve Bayes obteve a melhor acurácia, sendo o mais viável e eficaz. Este trabalho contribui com a literatura e gestão turística, entregando um modelo validado e replicável que permite automatizar a análise e transformar dados textuais não estruturados em inteligência competitiva para a melhoria do turismo.

Palavras-chave: Economia da Experiência. Inteligência Artificial. Aprendizado de Máquina. Predição. Classificação.

Abstract

This paper evaluates the application of Artificial Intelligence to the analysis of tourist reviews for enhancing tourism services. We tested several Machine Learning methods (Naive Bayes, Neural Networks, SVM, and LSTM) to classify and predict experiences from TripAdvisor reviews concerning the Jalapão and Superagui parks. The Naive Bayes method achieved the highest accuracy, demonstrating it is the most viable and effective method. This research contributes to both academic literature and tourism management by presenting a validated, replicable model that automates the analysis of unstructured text, transforming it into competitive intelligence to improve tourism.

Keywords: Experience Economy. Artificial Intelligence. Machine Learning. Prediction. Classification.

Introdução

O turismo é uma atividade importante para o desenvolvimento de uma região, e, viagens mais sustentáveis, roteiros flexíveis, busca por economia, consulta às redes sociais e apoio de ferramentas de Inteligência Artificial (IA) no planejamento de viagens, são algumas das tendências que estão no cenário turístico brasileiro, conforme destaca o Ministério do Turismo (2024). A relevância da Internet no contexto do turismo tem sido vigorosa, como pode ser visto nos estudos de Law et al. (2014), Boaria e Santos (2018) e Bandeira e De Menezes (2022), os quais aludem que os indivíduos, ao decidirem viajar, utilizam comentários das redes sociais tanto no que tange ao atrativo a ser visitado como em relação às opções da infraestrutura turística (alimentação e meios de hospedagem). Nesse sentido, torna-se pertinente aos gestores monitorar a reputação online dos empreendimentos com o fito de obter dados confiáveis para aperfeiçoar a prestação dos serviços e acurar a competitividade dos estabelecimentos hoteleiros (Bandeira & De Menezes, 2022), aspecto que pode ser extrapolado para os demais elementos do sistema turístico.

Diante do exposto, por ser o turismo uma atividade que necessita estar sempre em evolução, torna-se relevante entender as experiências relatadas pelos visitantes de uma localidade turística. A partir da compreensão das experiências explanadas, podem ser formuladas estratégias empresariais ou políticas públicas para a melhoria do turismo e, conseqüentemente, propiciar o desenvolvimento para a região (Kaizer et al., 2021).

A Internet, com notícias, sítios eletrônicos, postagens em mídias sociais ou fóruns de discussão, tem sido utilizada como fonte de dados para diversas avaliações. As pesquisas sobre sentimentos, por exemplo, envolvem um conjunto de técnicas e algoritmos concebidos para o tratamento de textos, reunindo conceitos de Aprendizado de Máquina (AM), Processamento de Linguagem Natural (PLN), mineração de dados, linguística e análise textual, tudo para interpretar fragmentos de textos e extrair o sentimento do escritor em relação a algum tema (Liu, 2012; Pang & Lee, 2008). O AM é um ramo da IA em que algoritmos possibilitam que computadores aprendam coisas novas e aperfeiçoadas a cada tarefa executada, permitindo, por exemplo, a classificação de imagens, a descoberta de novos medicamentos, a identificação de potenciais consumidores e até a composição de músicas (Davenport, 2018; Rose, 2020; Yang et al., 2013).

Estudos alinhados com a proposta de classificar a experiência dos turistas já foram executados por autores como Caracristi et al. (2021), Kaizer et al. (2021), Feger et al. (2022), (2023) e (2024), os quais se utilizam de comentários escritos no *website* de viagens *TripAdvisor* para identificar manualmente qual foi a experiência de turistas no Jalapão e no litoral de São Paulo. Os autores efetivamente encontraram os domínios de experiência relevantes para os turistas do Parque Estadual do Jalapão. Todavia, a análise manual de comentários de *websites* de viagem pode ser lenta e custosa, constituindo-se numa limitação para a aplicação da metodologia em outros atrativos turísticos.

Nesse sentido, a classificação automatizada de texto tem sido considerada vital para gerenciar e processar uma grande quantidade de documentos em formato digital que são difundidos na Internet. A maioria dos dados para classificação são coletados através de grupos de discussão, boletins e notícias transmitidas ou impressas (Kennedy et al., 2015), constituindo-se na possibilidade de acesso a um grande volume de informações. A partir do tratamento dos dados obtidos, pode-se disponibilizar conhecimentos que subsidiem decisões no campo do turismo e o emprego da IA abre espaço para aprofundamentos neste sentido.

Ghesh et al. (2024) apontam que ainda há possibilidade para pesquisas que abordem a identificação de ferramentas de IA mais eficazes para monitorar, medir e otimizar a experiência do usuário de serviços turísticos. Kim et al. (2024) e Al-Hyari et al. (2024) asseveram que a IA pode fornecer feedback imediato, mostrando as reações dos clientes a partir de diferentes abordagens e canais de trocas de mensagens e

impressões (comentários), contribuindo para criar produtos turísticos que atendam melhor aos requisitos dos clientes.

A contribuição que se pretende com este artigo é a identificação de variáveis e dimensões para avaliar os componentes e a experiência holística geral, que é uma lacuna importante apontada por Kabadayi et al. (2019). Além disso, pretende-se avançar na oferta de ferramentas que auxiliem na análise preditiva e na automação avaliativa por meio de novos sistemas baseados em IA, para antecipar requisitos de recursos e tarefas de personalização de serviços. Nesse sentido, busca ampliar a possibilidade de incorporação de dados espaço temporais sistemáticos que continua a ser um desafio, visto que os métodos disponíveis tratam as dimensões espaciais e temporais separadamente (Kim et al., 2024). A integração melhorada com a ajuda da IA oferece insights mais profundos relacionados com o comportamento dos turistas, fornecendo metodologias testadas para analisar dados textuais.

A comparação de desempenho dos métodos empregados é uma das tarefas mais críticas em AM (Jurman et al., 2012; Witten et al., 2017). A avaliação e validação do desempenho do método utilizado permite atestar o nível que o algoritmo pode performar, em um contexto real. Assim, além da contribuição relacionada com o método, a presente pesquisa contribui para avançar em estudos aplicando a IA em contextos distintos, sendo que o segmento de parques temáticos tem pouca aplicação de métodos utilizando esta tecnologia.

Este trabalho objetiva contribuir também ao avaliar e validar as previsões e classificações da experiência turística nos parques do Jalapão e do Superagui, aplicando métodos de AM em comentários publicados no *TripAdvisor*. É uma aplicação automatizada para a avaliação dos domínios da experiência (Pine & Gilmore, 1999) considerando um grande volume de dados. Relativo aos termos conceituais, considera-se a “classificação” como a eficiência dos modelos durante os treinamentos e testes, e a “predição” como a eficiência dos modelos treinados para previsões sobre dados novos, equivalente ao contexto real de aplicação. Assim, numa primeira etapa, é realizada a avaliação e a validação dos domínios da experiência para o parque Superagui, preditos por modelos gerados a partir dos dados do Jalapão. Na segunda etapa é avaliada a classificação de uma base unificada dos dados do Jalapão e Superagui.

A partir desta introdução, o restante do artigo contempla a revisão da literatura relacionada, seguida da apresentação dos procedimentos metodológicos, dos resultados, e culmina com as considerações finais.

Literatura Relacionada

Beni (2019) destaca que o turismo é uma atividade advinda de escolhas realizadas por diferentes motivações, que a envolve como, por exemplo, hospedagem, alimentação, lazer e entretenimento. Cita ainda que o consumidor busca por novas experiências em suas viagens e, segundo Pine e Gilmore (1999), uma experiência turística pode ser classificada dentro de quatro domínios: aprendizagem, entretenimento, estética e evasão. Estes domínios refletem o nível e a forma de envolvimento do indivíduo com o atrativo turístico, o qual deve proporcionar uma vivência nas quatro dimensões para que possibilite uma experiência memorável.

O setor de turismo tem mudado seus paradigmas através do avanço da tecnologia, com impactos que vão desde a organização da atividade até o comportamento dos consumidores (Cooper et al., 2007). Novas motivações de viagens, demandas e expectativas desses consumidores ensejam que, cada vez mais, roteiros e produtos turísticos se adaptem às suas necessidades e preferências.

De uma maneira geral, os comentários constantes das plataformas digitais e mídias sociais podem influenciar no planejamento de viagens. Souza e Machado (2017) apresentam uma análise sobre o grau dessa influência no planejamento e identificam o nível de confiança das informações dessas ferramentas. Na obra de Bandeira e De Menezes (2022), é apresentada uma análise de experiências dos hóspedes de hotéis de João Pessoa-PB, extraídas dos comentários do site *booking.com*.

Silva et al. (2021) analisaram as dimensões da experiência do cliente em parques temáticos, com base nas avaliações de viagens online. Mondo et al. (2020) efetuaram análise de serviços turísticos na Fortaleza de São José da Ponta Grossa, em Florianópolis-SC, utilizando o Tourqual, um modelo específico para avaliação da qualidade de serviços em atrativos turísticos.

Diversos trabalhos abordaram a análise de experiências turística no Parque Estadual do Jalapão (Caracristi et al., 2020; Caracristi et al., 2021; Feger et al., 2024a; Feger et al., 2024b; Kaizer et al., 2021; Marynowski et al., 2023). Esses trabalhos apresentaram análises realizadas a partir de comentários de redes sociais, com o objetivo de caracterizar a demanda para o referido destino turístico e identificar a qualidade da experiência percebida pelo turista.

Com base nas motivações do cliente, várias são as possibilidades de uso da IA e de novas tecnologias na atividade turística (Paes-Cesário & Melo, 2023). Crescencio et al. (2020) realizaram a classificação de sentimentos, positivos e negativos, a partir de comentários sobre a *Oktoberfest* Blumenau-SC, com os métodos Multinomial *Bayes* e *Word2Vec*, com acurácia e precisão superiores a 90%.

Relativo à organização do conteúdo de avaliações, ao processamento automático de grande número de informações e à extração de conhecimento, destacam-se a classificação e o ranqueamento de dicas de viagem, desenvolvido por Guy et al. (2017), com os métodos SVM e *Logistic Regression*. Destaca-se, também, o trabalho de McKenzie e Adams (2018) que explorou como as semelhanças e diferenças entre cidades e atrativos podem contribuir para o comportamento de turistas, com o uso do modelo probabilístico generativo *Latent Dirichlet Allocation* (LDA). Maschio e Marynowski (2022) e Marynowski et al. (2023) também aplicaram métodos de AM para classificação de comentários sobre experiências turísticas no Jalapão. Magalhães et al. (2021) apontam que os métodos automáticos, embora possam agilizar a organização das informações, carecem de validação com apoio de especialistas.

Acerca da validação de métodos de AM por especialistas, a qual ocorre com comparação entre a classificação automática e a manual, destaca-se a classificação de atrações em perfis turísticos desenvolvida por Magalhães et al. (2021). Essa obra compara os métodos SVM, Bayes, *Logistic Regression* e *Random Forest*, com a classificação realizada por especialistas da área de turismo. Manz et al. (2020) realizam a validação de um modelo de AM para a predição de risco de mortalidade de pacientes com câncer, a qual foi comparada com índices prognósticos validados na prática médica. Correia et al. (2021) validaram um modelo de AM para predição diagnóstica de doença coronariana, comparando-o com um modelo estatístico tradicional de regressão logística. Já Tsur et al. (2020) desenvolveram e validaram um modelo de AM para a predição de distopia de ombro, uma emergência obstétrica que pode ocorrer durante o parto, comparando-o com outras práticas médicas comumente usadas. Esses trabalhos caracterizam a importância do processo de validação das predições obtidas por modelos de AM.

Os atrativos turísticos, utilizados neste trabalho, possuem as seguintes características: o Parque Estadual do Jalapão, com cerca de 158 mil hectares, está localizado no município de Mateiros – TO, e é formado por rios, lagoas, cachoeiras, corredeiras, ressurgências hídricas (“fervedouros”) e dunas. Foi criado em 2001, para preservação dos recursos naturais da região e serve de abrigo para inúmeras espécies da flora, como o capim-dourado e o buriti (TOCANTINS, 2021). O Parque Nacional de Superagui, com cerca 33 mil hectares, é formado por reserva de Mata Atlântica, com extensa faixa de manguezais, dunas, lagoas, praias intocadas e florestas tropicais, e está localizado no município de Guaraqueçaba - PR. Foi criado em 1989, para proteger a flora e a fauna locais, e promover a pesquisa científica e o ecoturismo responsável (Instituto Chico Mendes de Conservação da Biodiversidade [ICMBio], 2024).

Procedimentos metodológicos

Para alcançar os objetivos propostos para este trabalho tomou-se uma metodologia mista, ou seja, quali-quantitativa (qualitativa e quantitativa), conforme definida por Pereira et al. (2018). Ao tratar dos dados

textuais, foi adotada uma abordagem qualitativa para relacionar aspectos advindos de opiniões de turistas, modelados nos domínios da experiência, relativos aos parques do Jalapão e Superagui, os quais podem ser generalizados para outros elementos turísticos. A escolha dessa perspectiva analítica é justificada pela grande relevância do aspecto subjetivo para a compreensão das estruturas sociais e organizacionais. A abordagem quantitativa foi empregada para medir o nível de acurácia obtido pelos modelos nas previsões e classificações nos domínios da experiência turística, a qual foi estabelecida por meio de indicadores numéricos. Essa mensuração permite estabelecer a conexão entre a observação empírica e a expressão matemática das relações quantitativas.

Os textos analisados neste trabalho advêm de comentários publicados no site *TripAdvisor* (tripadvisor.com.br) por turistas que visitaram o Parque do Jalapão entre os anos de 2011 e 2019, e o Parque do Superagui entre os anos de 2013 e 2022. Os comentários dos turistas foram combinados em grupos de quatro palavras (tetragramas ou quadrigramas). Segundo Caracristi et al. (2020), a partir de quadrigramas há uma melhor compreensão do sentido das palavras, constituindo aproximações de frases mais aptas a exprimir os níveis de experiência turística. Assim, a base de dados utilizada para as classificações e previsões são os quadrigramas mais frequentes, identificados a partir dos comentários dos turistas.

O processo de avaliação e validação da aplicação de AM para avaliação da experiência turística, proposto neste trabalho, é organizado em duas etapas.

Na 1ª Etapa é realizada a avaliação e a validação dos domínios da experiência para o Superagui, preditos por modelos gerados pelos métodos Redes Neurais Artificiais (modelo RNA), implementado por Maschio e Marynowski (2022), e *Gaussian Naïve Bayes* (modelo *Bayes*), *Support Vector Machine* (modelo SVM) e *Long Short-Term Memory Bidirecional* (modelo Bi-LSTM), implementados por Marynowski et al. (2023). Nesta etapa foram utilizados 5.114 quadrigramas do Superagui para a avaliação das previsões. Na Tabela 1 são mostrados alguns dos quadrigramas do Superagui, com a experiência predita por alguns dos modelos anteriormente implementados e classificados por especialistas do Curso de Turismo da Universidade Federal do Paraná (UFPR).

Tabela 1 - Exemplos de Quadrigramas Preditos e Classificados nos Domínios da Experiência

Quadrigrama	Modelos Preditivos			Classificador
	RNA	BAYES	Bi-LSTM	Especialista
trilha levar praia deserta	Aprendizagem	Entretenimento	Aprendizagem	Entretenimento
calma ideal quer paz	Entretenimento	Entretenimento	Evasão	Evasão
melhor sair bem cedo	Estética	Entretenimento	Evasão	Entretenimento
dia alugar bike trilha	Evasão	Entretenimento	Aprendizagem	Entretenimento
sol vale pena caminhada	Aprendizagem	Estética	Evasão	Entretenimento

Fonte: Elaborada pelos autores (2025).

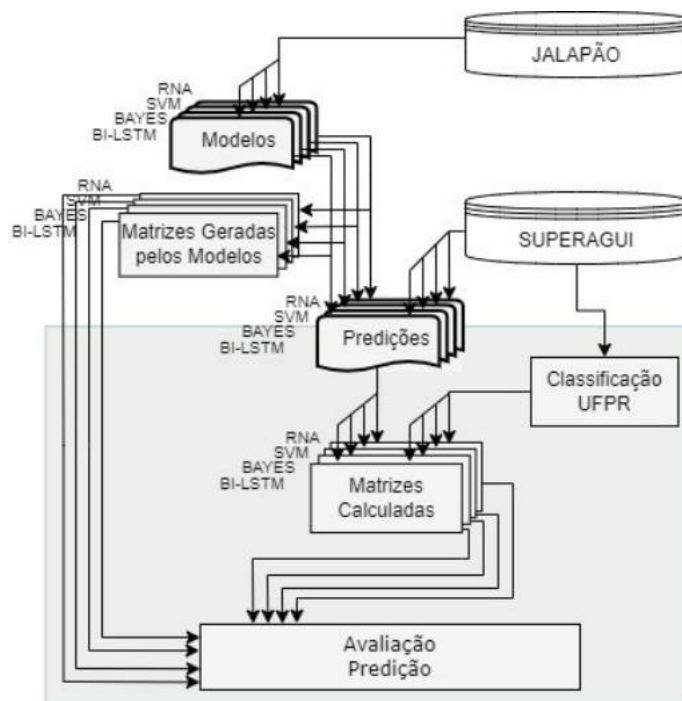
Na 2ª Etapa é realizada a avaliação dos domínios da experiência para uma base unificada Jalapão/Superagui, classificados pelos modelos RNA e *Bayes*. Nesta etapa foram utilizados os 5.114 quadrigramas da 1ª Etapa, somados a 1.256 quadrigramas do Jalapão.

A seguir são detalhadas as execuções de cada etapa deste trabalho.

1ª Etapa - Avaliação e validação das previsões

Está implementada conforme o fluxograma ilustrado na Figura 1, onde a área em destaque representa o segmento realizado neste trabalho. Esta etapa envolve as seguintes tarefas: Preparação dos dados, Cálculo das Matrizes de Confusão e Avaliação da Predição.

Figura 1 - Fluxograma da 1ª Etapa



Fonte: Elaborada pelos autores (2025).

Preparação dos dados

Durante a análise dos especialistas da UFPR, 82 quadrigramas não foram classificados nos domínios da experiência e 01 quadrigrama foi excluído por ser igual com outro da mesma base de dados. Diante disso, dos 5.114 quadrigramas do Superagui restaram 5.031, para processamento da predição.

Cálculo das Matrizes de Confusão

A matriz de confusão quantifica as amostras classificadas, correta e incorretamente pelo modelo, e é de onde se extrai a maioria das métricas de avaliação. O VP representa o verdadeiro positivo, i.e., o valor classificado corretamente como positivo. VN é o verdadeiro negativo, i.e., o valor classificado corretamente como negativo. FP é o falso positivo, i.e., o valor é classificado como positivo, mas ele é negativo. E o FN é o falso negativo, quando o valor é classificado como negativo, mas ele é positivo. Neste trabalho, é usada uma estrutura de matriz de confusão em que são considerados os valores preditos e os valores reais de cada classe.

Nesta 1ª Etapa, uma tabela com 5.031 quadrigramas classificados pelos especialistas da UFPR, dentro das quatro classes dos domínios de experiência, é comparada com as experiências preditas pelos modelos Bayes, RNA, SVM e Bi-LSTM. Dessas comparações são geradas as Matrizes de Confusão para cada um dos modelos utilizados.

Avaliação da Predição

As métricas de desempenho são utilizadas no AM em problemas de classificação, regressão ou clusterização. Na classificação de textos, as métricas de desempenho tornam possível averiguar o quanto o modelo é capaz de caracterizar um novo exemplo, quando lhe é apresentado (Liu, 2012). Elas auxiliam

tanto na avaliação do modelo, quanto na determinação daquele que seja mais adequado ao problema analisado. As principais métricas para avaliação de um modelo de classificação são:

Acurácia (*Accuracy*). Métrica mais comumente utilizada, expressa o quanto um modelo classificou corretamente uma amostra (Kulkarni et al., 2020). Ela é calculada através da equação (1):

$$accuracy = \frac{VP+VN}{(VP+VN+FP+FN)} \quad (1)$$

Precisão (*Precision*). Representa proporção de previsões corretas para uma categoria em relação ao total de previsões para uma categoria (De Castro & Ferrari, 2017). É calculada pela equação (2):

$$precision = \frac{VP}{VP+FP} \quad (2)$$

Revocação (*Recall*). Representa a proporção de previsões corretas para uma categoria em relação ao total de elementos de uma categoria (De Castro & Ferrari, 2017), calculada pela equação (3):

$$recall = \frac{VP}{VP+FN} \quad (3)$$

Medida F (*F1-score*). Representa a média harmônica entre a Precisão e a Revocação, calculada pela equação (4):

$$F1 - score = 2 \cdot \frac{(precision \cdot recall)}{(precision + recall)} \quad (4)$$

Para problemas de classificação multiclasse, a avaliação de desempenho pode ser calculada das seguintes maneiras (Sokolova & Lapalme, 2009):

Micro-média (*Micro Avg*). Calculada para cada uma das classes no conjunto $C=c_1, c_2, \dots, c_n$, e o valor cumulativo é utilizado para o cálculo da métrica para o modelo.

Macro-média (*Macro Avg*). É uma média aritmética das métricas calculadas para cada categoria, considerando igual peso para todas as categorias, independentemente do número de amostras de cada uma.

Média ponderada (*Weighted Avg*). As métricas de desempenho são calculadas para cada categoria e, ao final, a média é calculada ponderando os valores pelo suporte, isto é, pelo número de amostras de cada categoria.

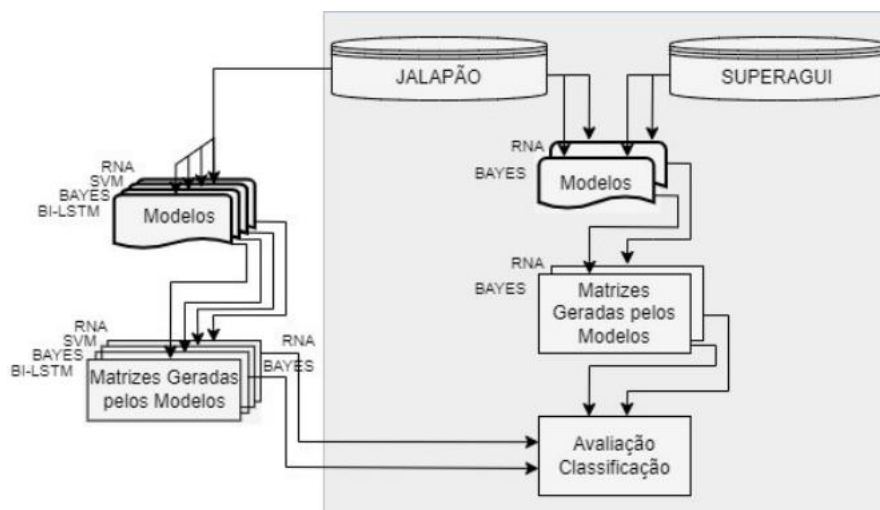
As Métricas de Desempenho foram calculadas, nesta etapa, a partir dos valores da Matriz de Confusão, com os recursos e funções do editor de planilhas Microsoft Excel.

Para a análise dos desempenhos dos modelos foram utilizadas, ainda, as medidas estatísticas de Média, Variância e Desvio-Padrão.

2ª Etapa - Avaliação da classificação de uma base unificada dos quadrigramas do Jalapão e Superagui

Esta tarefa é efetuada através dos modelos RNA e *Bayes*, e implementada conforme o fluxograma ilustrado na Figura 2. A área em destaque representa o segmento que foi realizado neste trabalho.

Figura 2 - Fluxograma da 2ª Etapa



Fonte: Elaborada pelos autores (2025)

O elevado tempo de processamento dos modelos BI-LSTM e SVM inviabilizou suas avaliações na classificação. Os processamentos excederam a capacidade do computador pessoal utilizado e ultrapassaram o limite de 12 horas disponibilizado pela versão gratuita do *Google Colab*.

Esta 2ª Etapa é composta por 4 fases: Preparação dos dados; Preparação do treinamento; Treinamento e Avaliação da Classificação.

Preparação dos dados

Nesta etapa foi realizada a estruturação da base de dados com o objetivo de condicioná-la ao treinamento e minimizar erros ou resultados imprecisos.

A base de dados unificada Jalapão/Superagui é obtida com a unificação dos 1.256 quadrigramas do Jalapão com 5.031 do Superagui, classificados pelos especialistas da UFPR, conforme discriminado na 1ª Etapa. Totalizaram 6.287 quadrigramas.

Foram excluídos 114 quadrigramas que apresentaram igualdade e pertenciam a classes diferentes, os quais foram unificados pelos Especialistas, restando 6.173 para processamento. A Tabela 2 apresenta a distribuição das quantidades dos quadrigramas nos quatro domínios da experiência, com destaque ao desbalanceamento da base com 4.262 para Entretenimento e 299 para Evasão.

Tabela 2 - Composição da Base Unificada Jalapão/Superagui por Domínio da Experiência

Domínio da Experiência	Quantidade
Entretenimento	4.262
Aprendizagem	838
Estética	774
Evasão	299
TOTAL	6.173

Fonte: Os autores (2025)

Preparação do treinamento

Nesta fase os quadrigramas passaram por um processamento para extrair suas características estruturais. Foi utilizada a função *unidecode* para converter os quadrigramas em uma versão sem acentos e sem cedilhas.

Para a vetorização dos quadrigramas, foi usada a função *tokenizer*, da biblioteca *Keras* versão 2.13.1, com o modelo matemático *word embeddings*, para a implementação do método RNA. Para o método *Bayes* a vetorização foi realizada com a medida TF-IDF (*Term Frequency–Inverse Document Frequency*), por meio da função *TfidfVectorizer*, da biblioteca *Scikit-learn* versão 1.2.2.

Devido ao desbalanceamento das classes, evidenciado na Tabela 2, o método Bayes utiliza a técnica SMOTE (*Synthetic Minority Oversampling Technique*), o qual cria exemplos da classe minoritária por meio da interpolação de diversos exemplos que se encontram próximos (Chawla, 2002).

Os modelos gerados para classificação da base Jalapão/Superagui seguiram as mesmas configurações da classificação da base do Jalapão, implementados por Maschio e Marynowski (2022) e Marynowski et al. (2023). Isso justifica a ausência do algoritmo de balanceamento (SMOTE) no método RNA.

Treinamento

Enquanto um parâmetro é uma variável de configuração interna ao modelo de aprendizagem, cujo valor pode ser estimado a partir dos dados de treinamento (Claesen & Moor, 2015), os hiperparâmetros são configurações usadas para controlar o comportamento do algoritmo e definir a estrutura do modelo (Goodfellow et al., 2016).

O modelo RNA foi implementado com os seguintes hiperparâmetros: tamanho da camada de *embedding*: 200; função de ativação: *Softmax*; função de otimização: *Adam*; tamanho de lote (*batch size*): 40 e número de épocas (*epochs*): 200. A divisão de treinamento e teste foi de 85 e 15, respectivamente.

O modelo *Bayes* foi implementado com validação cruzada *K-Fold* ($k = 50$), parâmetro *smoothing* de 10^{-4} e divisão de treinamento e teste de 67 e 33, respectivamente.

Avaliação da Predição

Nesta fase, as mesmas Métricas de Desempenho discriminadas na 1ª Etapa foram geradas automaticamente pelos modelos preditivos utilizados na classificação (RNA e *Bayes*), através da função *classification_report* da biblioteca *Scikit-learn* versão 1.2.2.

Apresentação e análise dos resultados

Para facilitar a análise e interpretação dos resultados, eles são apresentados, primeiramente, em relação às predições realizadas (1ª Etapa da metodologia desta pesquisa) e, posteriormente, em relação à geração de novos modelos através de classificações da base unificada (2ª Etapa). Os resultados estão sistematizados nas tabelas apresentadas, onde os números mais significativos estão em negrito.

Dentre as Métricas de Desempenho, a acurácia é a que expressa o quão corretamente um modelo classificou uma amostra. Optou-se, neste trabalho, pelo uso da acurácia como a métrica principal de avaliação pela sua praticidade, simplicidade e facilidade de entendimento pelo analista do segmento turístico.

Considerando que nas classificações e predições, nos domínios da experiência, as métricas de precisão e *recall* possuem o mesmo nível de importância, foi considerado o *f1-score*, média harmônica entre elas, como a métrica secundária para avaliação de desempenho dos modelos.

Predições da experiência para Superagui

Acompanhando o trabalho de Magalhães et al. (2021), essa seção reforça a necessidade de validação dos resultados obtidos pelos modelos, comparando as suas classificações com a classificação realizada por especialistas da área de turismo. O objetivo é validar as métricas obtidas por meios automatizados com os

obtidos manualmente, a fim de garantir a equivalência dos resultados e proporcionar aos gestores uma avaliação da experiência considerando um grande volume de dados, conforme indicado por Ghesh et al. (2024).

Os modelos gerados para o Jalapão são utilizados para a predição da experiência para o Superagui. Os resultados são apresentados através das matrizes de confusão dos modelos, calculadas manualmente, e das suas respectivas métricas de desempenho. As métricas são apresentadas, nas tabelas, em formato similar àquele gerado pelas bibliotecas utilizadas, com as classes (domínios da experiência), as métricas de cada classe (*precisão*, *recall*, *f1-score*) e o número de amostras por classe (*support*) exibidas na parte superior e, na parte inferior da tabela, estão as métricas globais do modelo (*acurácia*, *micro avg*, *macro avg* e *weighted avg*). Foram utilizadas as seguintes abreviações: Aprendizagem (AP), Entretenimento (EN), Estética (ES) e Evasão (EV).

A Tabela 3 apresenta a Matriz de Confusão do modelo *Bayes*, dentro dos domínios da experiência. Observa-se que a classe Entretenimento foi a que obteve o maior número de VP, com 3098 amostras preditas corretamente, de um total de 3786 amostras. Destaca-se que na classe Evasão, com 221 amostras, nenhuma foi predita corretamente, sendo a maioria delas, 174 amostras, preditas incorretamente como Entretenimento.

Tabela 3 - Matriz de Confusão do Modelo Bayes na Predição da Experiência para Superagui

Classe Real	Classe Predita			
	Aprendizagem	Entretenimento	Estética	Evasão
Aprendizagem	155	327	56	0
Entretenimento	392	3098	296	0
Estética	26	167	293	0
Evasão	15	174	32	0

Fonte: Elaborada pelos autores (2025).

A Tabela 4 apresenta as Métricas de Desempenho para o modelo *Bayes*. Verifica-se que o modelo alcançou nas predições uma acurácia de 70,48% e, considerando a média ponderada (*weighted avg*), um valor de *f1-score* de 68,42%.

Tabela 4 - Métricas de Desempenho do Modelo Bayes na Predição da Experiência para Superagui

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Aprendizagem	26,36%	28,81%	27,53%	538
Entretenimento	82,26%	81,83%	82,04%	3786
Estética	43,28%	60,29%	50,39%	448
Evasão	0,00%	0,00%	0,00%	221
Métricas Globais				
<i>Accuracy</i>			70,48%	5031
<i>Micro Avg</i>	70,48%	70,48%	70,48%	5031
<i>Macro Avg</i>	37,98%	42,73%	39,99%	5031
<i>Weighted Avg</i>	67,79%	69,34%	68,42%	5031

Nota. Os resultados mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 5 apresenta a Matriz de Confusão do modelo gerado pelo modelo SVM. Observa-se que a classe Entretenimento foi a que obteve o maior número de VP, com 3325 amostras preditas corretamente pelo

modelo. A classe Evasão, com 221 amostras, só obteve 11 amostras preditas corretamente e, a grande maioria delas, predita erroneamente como Entretenimento.

Tabela 5 - Matriz de Confusão do Modelo SVM na Predição da Experiência para Superagui

Classe Real	Classe Predita			
	Aprendizagem	Entretenimento	Estética	Evasão
Aprendizagem	149	369	18	2
Entretenimento	358	3325	99	4
Estética	32	286	162	6
Evasão	19	174	17	11

Nota. O maior valor de VP está destacado em negrito.

Fonte: Elaborada pelos autores (2025).

O resultado verificado nas Tabelas 4 e 5 são coerentes com os encontrados nos estudos de Caracristi et al. (2021) e Kaizer et al. (2021), no que concerne ao domínio Entretenimento ser mais evidenciado, tendo o maior número de acertos (VP), e o domínio Evasão foi a que teve menor volume de tetragramas identificados.

A Tabela 6 mostra as Métricas de Desempenho do modelo SVM. Verifica-se que o modelo alcançou, nas predições, uma acurácia de 72,49%, e *f1-score* de 69,19%.

Tabela 6 - Métricas de Desempenho do Modelo SVM na Predição da Experiência para Superagui

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Aprendizagem	26,70%	27,70%	27,19%	538
Entretenimento	80,04%	87,82%	83,75%	3786
Estética	54,73%	33,33%	41,43%	448
Evasão	47,83%	4,98%	9,02%	221
Métricas Globais				
<i>Accuracy</i>			72,49%	5031
<i>Micro Arg</i>	72,49%	72,49%	72,49%	5031
<i>Macro Arg</i>	52,33%	38,46%	40,35%	5031
<i>Weighted Arg</i>	69,33%	71,31%	69,19%	5031

Nota. Os resultados mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 7 mostra a Matriz de Confusão do modelo Bi-LSTM. Observa-se que na classe Entretenimento, 3495 foram incorretamente previstas como Aprendizagem pelo modelo. A classe Aprendizagem foi a que obteve o maior número de VP, com 510 preditas corretamente, das 538 existentes. A classe Evasão teve o menor número de VP, com 19 amostras preditas corretamente.

Tabela 7 - Matriz de Confusão do Modelo Bi-LSTM na Predição da Experiência para Superagui

Classe Real	Classe Predita			
	Aprendizagem	Entretenimento	Estética	Evasão
Aprendizagem	510	10	2	16
Entretenimento	3495	195	26	70
Estética	384	17	72	3

Evasão	185	15	2	19
--------	-----	----	---	----

Nota. Os maiores valores de VP estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 8 apresenta as Métricas de Desempenho do modelo Bi-LSTM. Observa-se que esse modelo obteve uma acurácia de 15,82%, implicando em uma Taxa de Erro acima de 84%, além de *f1-score* de 12,09%.

Tabela 8 - Métricas de Desempenho do Modelo Bi-LSTM na Predição da Experiência para Superagui

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Aprendizagem	11,15%	94,80%	19,95%	538
Entretenimento	82,28%	5,15%	9,69%	3786
Estética	70,59%	14,81%	24,49%	448
Evasão	16,10%	8,60%	11,21%	221
Métricas Globais				
<i>Accuracy</i>			15,82%	5031
<i>Micro Avg</i>	15,82%	15,82%	15,82%	5031
<i>Macro Avg</i>	45,03%	30,84%	16,34%	5031
<i>Weighted Avg</i>	69,49%	15,57%	12,09%	5031

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 9 apresenta a Matriz de Confusão do modelo RNA. Observa-se que da classe Entretenimento, 2381 amostras foram previstas corretamente, melhor resultado obtido pelo modelo. A classe Evasão teve o menor número de VP, com 7 amostras previstas corretamente, das 221 existentes.

Tabela 9 - Matriz de Confusão do Modelo RNA na Predição da Experiência para Superagui

Classe Real	Classe Predita			
	Aprendizagem	Entretenimento	Estética	Evasão
Aprendizagem	129	341	57	11
Entretenimento	934	2381	382	89
Estética	115	325	38	8
Evasão	49	147	18	7

Nota. O maior valor de VP está destacado em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 10 apresenta as Métricas de Desempenho do modelo RNA. Observa-se que o modelo obteve uma acurácia de 50,79% e *f1-score* de 52,96%.

Tabela 10 - Métricas de Desempenho do Modelo RNA na Predição da Experiência para Superagui

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Aprendizagem	10,51%	23,98%	14,62%	538
Entretenimento	74,55%	62,89%	68,22%	3786
Estética	7,68%	7,82%	7,75%	448
Evasão	6,09%	3,17%	4,17%	221
Métricas Globais				
<i>Accuracy</i>			50,79%	5031

<i>Micro Avg</i>	50,79%	50,79%	50,79%	5031
<i>Macro Avg</i>	24,71%	24,46%	23,69%	5031
<i>Weighted Avg</i>	57,29%	49,96%	52,96%	5031

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 11 apresenta um quadro comparativo do desempenho dos modelos nas predições com as métricas de acurácia, precisão, *recall* e o *f1-score*. A maior acurácia nas predições foi obtida pelo modelo SVM, com 72,49%, seguido por *Bayes*, com 70,48% e por RNA, com 50,79%. A menor acurácia foi obtida pelo modelo Bi-LSTM, com 15,82%. O *f1-score* foi maior pelo modelo SVM, com 40,35% e mais baixo pelo modelo Bi-LSTM, de 16,34%. Destaca-se o baixo desempenho do modelo Bi-LSTM, com cerca de 15% de acurácia, abaixo dos demais modelos nas predições e inferior ao seu próprio desempenho, nas classificações dos dados do Jalapão.

O baixo desempenho do modelo Bi-LSTM pode representar um caso em que o modelo esteja demasiadamente ajustado aos dados de treinamento (*overfitting*). Nessa situação, o modelo pode ter se especializado nos dados de treinamento para os dados do Jalapão, mas não conseguiu generalizar para o contexto ligeiramente diferente, que é o de Superagui.

Tabela 11 - Desempenho dos Modelos na Predição da Experiência para Superagui

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Bayes	70,48%	37,98%	42,73%	39,99%
SVM	72,49%	52,33%	38,46%	40,35%
Bi-LSTM	15,82%	45,03%	30,84%	16,34%
RNA	50,79%	24,71%	24,46%	23,69%

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

Verifica-se, portanto, que os modelos SVM e *Bayes* tiveram desempenhos próximos nas predições, perto dos 70% de acurácia. Ambos os modelos demonstram boa capacidade para executar as predições da experiência, a partir de quadrigramas de base de dados diferentes daquelas que foram treinados. Embora trabalhos como o de Manz et al. (2020) e Tsur et al. (2020) exijam uma acurácia maior, pois são voltados para predição médica, considera-se que os resultados dos modelos aqui apresentados já poderiam ser empregados para uma avaliação inicial da experiência em outros atrativos. Trata-se de uma tendência que pode ser ponderada pelo gestor e contribuir para avaliar a experiência holística geral do turista, conforme lacuna apontada por Kabadayi et al. (2019). Entretanto, é importante aprimorar a precisão da predição obtida e isso é abordado a seguir.

Classificação dos quadrigramas da base unificada Jalapão/Superagui

Para evoluir a eficácia na avaliação da experiência turística de forma otimizada e automatizada, a base de dados do Jalapão e do Superagui foi unificada, e novos modelos foram gerados e avaliados, buscando assim gerar indicadores mais precisos conforme indicado por Ghesh et al. (2024).

A Tabela 12 apresenta a Matriz de Confusão do modelo RNA para a classificação da base unificada. Observa-se que da classe Entretenimento, 516 amostras foram classificadas corretamente, melhor resultado obtido pelo modelo. A classe Evasão teve o menor número de VP, com 23 amostras preditas corretamente, das 52 existentes.

Tabela 12 - Matriz de Confusão do Modelo RNA na Classificação da Base Unificada

Classe Real	Classe Predita			
	Aprendizagem	Entretenimento	Estética	Evasão
Aprendizagem	58	65	12	1
Entretenimento	55	516	24	32
Estética	2	24	76	9
Evasão	0	23	6	23

Nota. O maior valor de VP está destacado em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 13 apresenta as Métricas de Desempenho do modelo RNA na classificação. Em que pese o processamento ter sido realizado com os dados desbalanceados, verifica-se que o modelo obteve uma acurácia de 72,68% e f1-score de 72,63%, considerando a média ponderada (*weighted avg*).

Tabela 13 - Métricas de Desempenho do Modelo RNA na Classificação da Base Unificada

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Aprendizagem	50,43%	42,65%	46,22%	136
Entretenimento	82,17%	82,30%	82,23%	627
Estética	64,41%	68,47%	66,38%	111
Evasão	35,38%	44,23%	39,32%	52
Métricas Globais				
<i>Accuracy</i>			72,68%	926
<i>Micro Avg</i>	72,68%	72,68%	72,68%	926
<i>Macro Avg</i>	58,10%	59,41%	58,53%	926
<i>Weighted Avg</i>	72,75%	72,68%	72,63%	926

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 14 apresenta a Matriz de Confusão do modelo *Bayes* na classificação dos quadrigramas da base unificada. Observa-se que as classes foram balanceadas para 4.262 amostras. Embora o número de VP de todas as classes tenha sido alto, o maior número de VP foi da classe Evasão, com 4.238 amostras classificadas corretamente. Houve apenas 24 amostras com classificações incorretas, todas classificadas como Estética. A classe Entretenimento foi a que obteve o menor número de VP, com 2.293 amostras.

Tabela 14 - Matriz de Confusão do Modelo Bayes na Classificação da Base Unificada

Classe Real	Classe Predita			
	Aprendizagem	Entretenimento	Estética	Evasão
Aprendizagem	3589	104	396	173
Entretenimento	845	2293	606	518
Estética	35	11	3836	380
Evasão	0	0	24	4238

Nota. O maior valor de VP está destacado em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 15 apresenta as Métricas de Desempenho do modelo *Bayes* na classificação. Verifica-se que o modelo alcançou uma acurácia de 81,86% e *f1-score* de 80,90%.

Tabela 15 - Métricas de Desempenho do Modelo Bayes na Classificação da Base Unificada

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
Aprendizagem	80,31%	84,21%	82,21%	4262
Entretenimento	95,22%	53,80%	68,76%	4262
Estética	78,90%	90,00%	84,09%	4262
Evasão	79,83%	99,44%	88,56%	4262
Métricas Globais				
<i>Accuracy</i>			81,86%	17048
<i>Micro Avg</i>	81,86%	81,86%	81,86%	17048
<i>Macro Avg</i>	83,56%	81,86%	80,90%	17048
<i>Weighted Avg</i>	83,56%	81,86%	80,90%	17048

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

A Tabela 16 apresenta um quadro comparativo do desempenho dos modelos nas classificações com a métrica de acurácia e, considerando a macro-média, as métricas de precisão, *recall* e o *f1-score*. Verifica-se que o modelo *Bayes* obteve o melhor desempenho na classificação, com acurácia de 81,86% e *f1-score* de 80,90%. O modelo RNA obteve uma acurácia de 72,68% e *f1-score* de 58,53%. Destaca-se que ambos os modelos foram utilizados com ajuste de hiperparâmetros. Ainda, o modelo *Bayes* envolveu o SMOTE para balanceamento dos dados.

Tabela 16 - Desempenho dos modelos na Classificação da Base Unificada

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Bayes	81,86%	83,56%	81,86%	80,90%
RNA	72,68%	58,10%	59,41%	58,53%

Nota. Os maiores valores estão destacados em negrito.

Fonte: Elaborada pelos autores (2025).

Avaliação de desempenho geral dos modelos nas classificações e previsões

As Tabelas 17, 18, 19 e 20 sintetizam o desempenho dos modelos nas classificações executadas com os dados do Jalapão, implementadas por Maschio e Marynowski (2022) e Marynowski et al. (2023), bem como nas previsões avaliadas para o Superagui e nas classificações executadas para a base unificada Jalapão/Superagui.

Além das métricas já consideradas de acurácia, precisão, *recall* e *f1-score*, também são utilizadas a média, a variância e o desvio-padrão das execuções realizadas por cada modelo. Cabe destacar que estes quadros-resumo objetivam demonstrar o comportamento de cada modelo no decorrer das tarefas realizadas. Destaca-se que, devido aos modelos SVM e Bi-LSTM não terem executado a classificação da base unificada, por insuficiência de recursos para seus processamentos, uma comparação entre os modelos merece maior atenção e uma adequação na sua análise.

A Tabela 17 apresenta a acurácia dos modelos nas tarefas de classificação do Jalapão, previsão do Superagui e classificação da base unificada Jalapão/Superagui.

Observa-se que na comparação dos modelos, nas classificações dos quadrigramas do Jalapão e na previsão do Superagui, todos tiveram uma redução no valor da acurácia. O modelo *Bayes* obteve a menor redução de acurácia, baixou de 81,04% na classificação para 70,48% na previsão dos dados do Superagui, seguido pelo modelo SVM que reduziu de 88,14% para 72,49%, com uma pequena variância de 1,22%.

A

redução mais acentuada ocorreu no modelo Bi-LSTM, que caiu de 83,65% na classificação para 15,82% na predição, com grande dispersão, tendo uma variância de 23%, seguida pelo modelo RNA, que caiu de 77,78% para 50,79%.

Considerando as duas classificações, dos dados do Jalapão e dados unificados, verifica-se que os modelos *Bayes* e RNA tiveram pouca variação na acurácia. O modelo *Bayes* subiu de 81,04% para 81,86%, e o modelo RNA baixou de 77,78% para 72,68%, ambos mantiveram a constância nos dois processos de classificação. O modelo *Bayes* teve, também, um incremento de 0,82% na acurácia, na segunda classificação.

Tabela 17 - Acurácia dos Modelos nas Tarefas de Classificação e Predição

	Bayes	RNA	SVM	Bi-LSTM
Classificação Jalapão*	81,04%	77,78%	88,14%	83,65%
Predição Superagui	70,48%	50,79%	72,49%	15,82%
Classificação Unificada	81,86%	72,68%	-	-
Medidas estatísticas				
Média	77,80%	67,08%	80,31%	49,73%
Variância	0,40%	2,06%	1,22%	23,00%
Desvio-padrão	6,35%	14,34%	11,06%	47,96%

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025) e *Maschio e Marynowski (2022) e Marynowski et al. (2023).

A Tabela 18 apresenta a precisão dos modelos nas tarefas de classificação do Jalapão, predição do Superagui e classificação para a base unificada Jalapão/Superagui. Comparando a classificação Jalapão e a predição Superagui, verifica-se que todos os modelos tiveram uma redução acentuada, com destaque para os modelos RNA, com queda de 73,74% para 24,71% e *Bayes*, com queda de 81,82% para 37,98%. Os modelos SVM e Bi-LSTM tiveram suas medidas de dispersão próximas, com variâncias de 6,53% e 7,46% e desvios-padrões de 25,56% e 27,31%, respectivamente.

Considerando as duas classificações, dos dados do Jalapão e da base unificada Jalapão/Superagui, verifica-se que o modelo *Bayes* subiu sua precisão de 81,82% para 83,56%, e o RNA baixou a precisão de 73,74% para 58,10%.

Tabela 18 - Precisão dos Modelos nas Tarefas de Classificação e Predição

	Bayes	RNA	SVM	Bi-LSTM
Classificação Jalapão*	81,82%	73,74%	88,47%	83,65%
Predição Superagui	37,98%	24,71%	52,33%	45,03%
Classificação Unificada	83,56%	58,10%	-	-
Medidas estatísticas				
Média	67,79%	52,18%	70,40%	63,34%
Variância	6,67%	6,27%	6,53%	7,46%
Desvio-padrão	25,83%	25,05%	25,56%	27,31%

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025) e *Maschio e Marynowski (2022) e Marynowski et al. (2023).

A Tabela 19 apresenta a medida recall para os modelos nas tarefas de classificação do Jalapão, predição do Superagui e classificação da base unificada Jalapão/Superagui.

Considerando a classificação com os dados do Jalapão e a predição para Superagui, verifica-se que todos os modelos tiveram uma redução acentuada da classificação para a predição, com destaque para os modelos RNA, com queda de 72,20% para 24,46% e Bi-LSTM, com queda de 83,91% para 30,84%.

Considerando as duas classificações, do Jalapão e da base unificada Jalapão/Superagui, verifica-se que o modelo *Bayes* aumentou o recall de 81,04% para 81,86%, e o RNA reduziu de 72,20% para 59,41%.

Tabela 19 - Recall dos Modelos nas Tarefas de Classificação e Predição

	Bayes	RNA	SVM	Bi-LSTM
Classificação Jalapão*	81,04%	72,20%	88,14%	83,91%
Predição Superagui	42,73%	24,46%	38,46%	30,84%
Classificação Unificada	81,86%	59,41%	-	-
Medidas estatísticas				
Média	68,55%	52,02%	63,30%	57,37%
Variância	5,00%	6,11%	12,34%	14,08%
Desvio-padrão	22,36%	24,71%	35,13%	37,53%

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025) e *Maschio e Marynowski (2022) e Marynowski et al. (2023).

A Tabela 20 apresenta o *f1-score* dos modelos nas tarefas de classificação do Jalapão, predição do Superagui e classificação da base unificada Jalapão/Superagui.

Na comparação dos modelos na execução das classificações dos dados do Jalapão e na predição para Superagui, todos tiveram uma redução no valor do *f1-score*. Destaca-se que o modelo Bi-LSTM, teve redução de *f1-score* de 83,55% na classificação, para 16,34% na predição dos dados do Superagui, seguido pelo modelo RNA que reduziu essa mesma métrica de 72,87% para 23,69%.

Considerando as duas classificações, dos dados do Jalapão e dados unificados, verifica-se que o modelo *Bayes* subiu de 79,62% para 80,90%, e o RNA baixou de 72,87% para 58,53%.

Tabela 20 - F1-score dos Modelos nas Tarefas de Classificação e Predição.

	Bayes	RNA	SVM	Bi-LSTM
Classificação Jalapão*	79,62%	72,87%	88,24%	83,55%
Predição Superagui	39,99%	23,69%	40,35%	16,34%
Classificação Unificada	80,90%	58,53%	-	-
Medidas estatísticas				
Média	66,84%	51,70%	64,29%	49,94%
Variância	5,41%	6,40%	11,47%	22,59%
Desvio-padrão	23,26%	25,29%	33,86%	47,53%

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025) e *Maschio e Marynowski (2022) e Marynowski et al. (2023).

A Tabela 21 apresenta o desempenho geral dos modelos nas tarefas de classificação do Jalapão, predição do Superagui e classificação da base unificada Jalapão/Superagui.

De uma maneira geral, considerando a classificação dos dados do Jalapão e a predição dos dados do Superagui, o modelo SVM obteve as maiores métricas na classificação dos dados do Jalapão com 88,14% de acurácia, 88,47% de precisão, 88,14% de recall e 88,24% de *f1-score* e, com exceção do recall, obteve as maiores métricas na predição dos dados do Superagui com 72,49% de acurácia, 52,33% de precisão e

40,35% de *f1-score*. Na predição do Superagui o modelo *Bayes* obteve um desempenho muito próximo do SVM, com 70,48% de acurácia.

Na classificação dos dados unificados, verifica-se que o modelo *Bayes* obteve o melhor desempenho com 81,86% de acurácia, 83,56% de precisão, 81,86% de *recall* e 80,90% de *f1-score*. Esses valores foram superiores aos alcançados, pelo mesmo modelo, na classificação do Jalapão e predição do Superagui, o que demonstra que foi mais eficiente que suas versões anteriores.

Tabela 21 - Desempenho Geral dos Modelos nas Tarefas de Classificação e Predição.

	Bayes	RNA	SVM	Bi-LSTM
Classificação Jalapão*				
Acurácia	81,04%	77,78%	88,14%	83,65%
Precisão	81,82%	73,74%	88,47%	83,65%
<i>Recall</i>	81,04%	72,20%	88,14%	83,91%
<i>F1-score</i>	79,62%	72,87%	88,24%	83,55%
Predição Superagui				
Acurácia	70,48%	50,79%	72,49%	15,82%
Precisão	37,98%	24,71%	52,33%	45,03%
<i>Recall</i>	42,73%	24,46%	38,46%	30,84%
<i>F1-score</i>	39,99%	23,69%	40,35%	16,34%
Classificação Unificada				
Acurácia	81,86%	72,68%	-	-
Precisão	83,56%	58,10%	-	-
<i>Recall</i>	81,86%	59,41%	-	-
<i>F1-score</i>	80,90%	58,53%	-	-

Nota. Os valores mais relevantes estão destacados em negrito.

Fonte: Elaborada pelos autores (2025) e *Maschio e Marynowski (2022) e Marynowski et al. (2023).

Discussão dos Resultados

Como já discutido nas linhas anteriores, com o advento da Internet e do comportamento humano de compartilhar experiências, inclusive no que diz respeito a viagens, há um grande volume de dados que podem ser utilizados pelas organizações a fim de orientar suas estratégias mercadológicas, conforme Boaria e Santos (2018), Bandeira e De Menezes (2022) e Law et al., (2014). Por isso a proposta desta investigação de testar diferentes modelos a fim de verificar os que mais se adaptam para a interpretação de comentários, sob o prisma da teoria da economia da experiência (Pine & Gilmore, 1999), se mostrou relevante.

Inicialmente, contribuímos para reduzir a lacuna apontada por Kabadayi et al. (2019) na avaliação da experiência holística geral do turista, apresentando que os modelos Bayes e SVM foram identificados como os mais propícios para serem utilizados na predição da experiência dos turistas do Superagui, considerando os modelos gerados para o parque do Jalapão. Esses modelos foram capazes de atingir níveis de eficácia acima de 70% na classificação e predição. Maschio e Marynowski (2022) e Marynowski et al. (2023) apresentaram resultados acima de 80%, mas apenas considerando o mesmo conjunto de dados para treinamento e avaliação da predição. Esse aspecto reforça o alerta de Magalhães et al. (2021) quanto à necessidade de comparar as classificações elaboradas por meios automatizados com os manuais, a fim de identificar a acurácia "real" dos métodos. Ainda assim, infere-se que a acurácia de 70% na avaliação automatizada de outro atrativo se presta a suportar as empresas no sentido de criar indicadores de tendências que permitam monitorar, medir e otimizar a experiência de usuários de serviços turísticos por meio de uma ferramenta de IA, conforme apontado por Ghesh et al. (2024).

Ainda, embora numa proporção menor, observa-se que os domínios das experiências se apresentaram na mesma ordem de classificação obtida em Caracristi et al. (2021) e Kaizer et al. (2021). Confirma-se então que uma taxa de acerto de 70% possibilita ao gestor ponderar quanto à necessidade de atuação para melhorar o desempenho em determinados domínios da experiência, bem como, no caso de replicação da análise após a implementação de mudanças, para verificar se as medidas trouxeram os resultados esperados (Kaizer et al., 2021). Nesse sentido, a automação do processo é de grande valia, visto que demanda menos esforço, e pode ser replicado sistematicamente pela organização a fim de monitorar o desempenho, após implementação de medidas de melhoria no atrativo.

Entretanto, a indicação de tendência apresentada pela ferramenta pode levar a uma conclusão equivocada do gestor e deve ser interpretada com cautela. No caso aqui apresentado, o domínio Evasão foi o pior, com nenhum acerto, resultados assemelhados aos encontrados por Caracristi et al. (2021) e Kaizer et al. (2021), mas isso não indica sua inexistência. Podem significar que as experiências de Evasão sejam fracas e não gerem comentários distintos, já que as respostas envolvem a espontaneidade dos turistas, ou que as campanhas de marketing que promovem o parque, como um refúgio, não estejam alinhadas com a percepção do público. Deve-se considerar ainda o reportado por Kozinets (2014) ao debater o uso de dados disponibilizados na internet, especialmente comentários de usuários, que muitas vezes revelam os seus sentimentos, entretanto, não contemplando todas as variáveis previstas no referencial teórico utilizado pelo pesquisador. Destaca-se, portanto, que estes resultados devem ser utilizados com reservas em possíveis implicações gerenciais.

Acerca dos resultados da classificação, quando a base de dados foi unificada e novos modelos foram gerados, é possível verificar que os dois métodos reimplementados (Bayes e RNA) melhoraram todas as suas métricas de desempenho, quando comparados às suas predições. Tal fato é devido à formulação de novos modelos, treinados e testados com uma base de dados com maior número de amostras e a classificação automatizada, onde a métrica é obtida a partir de parte da base de dados usada para treinamento. Esse refinamento é importante e previsto no AM para melhorar a eficácia dos modelos, conforme indicado por Witten et al. (2017). Mas não indica a efetividade em caso de aplicação em um outro atrativo, sendo necessário uma avaliação crítica para uma nova aplicação.

Quanto ao desempenho dos métodos, cabe ponderar que o desempenho do RNA poderia ser diferente com a aplicação do SMOTE no balanceamento de dados. Entretanto, esta ação não foi realizada para manter a compatibilidade com os trabalhos de Maschio e Marynowski (2022) e Marynowski et al. (2023). Além disso, o balanceamento dos dados para a classificação da base unificada Jalapão/Superagui merece atenção e a realização de novas avaliações, principalmente para os modelos RNA, SVM e Bi-LSTM, que podem apresentar ganho de efetividade e não puderam ser realizados devido a limitação de recursos computacionais para execução. Sugere-se ainda, uma avaliação do desempenho dos modelos considerando outra métrica principal, como o F1-score, que pode ser mais representativo e robusto para classes desbalanceadas.

Considerações finais

O presente estudo se alinha às pesquisas que buscam auxiliar os gestores na tomada de decisões com base em dados disponíveis em redes sociais. No campo do turismo, contribui para a análise de grandes volumes de dados disponibilizados em redes sociais, conforme Boaria e Santos (2018), Bandeira e De Menezes (2022) e Law et al., (2014) relatam, e visto que, se tratados manualmente, demandam um grande esforço e inibem o seu uso sistemático, conforme indicado nos estudos de Caracristi et al. (2021), Kaizer et al. (2021) e Feger et al. (2022).

A realização deste trabalho contribui com a literatura acadêmica oferecendo a avaliação da aplicação de modelos de AM que, neste caso, pode representar um facilitador no processo decisório de gestores e subsidiar planejamentos futuros para a melhora da experiência turística dos visitantes do parque do

Jalapão e do Superagui. Essas ações podem trazer benefícios econômicos e sociais, além de contribuir para a proteção ambiental e a valorização das culturas locais.

Esses modelos possibilitam, também, iniciar o processo de automatização de avaliação da experiência em outros atrativos turísticos por meio da IA, conforme apontado pelo Ministério do Turismo (2024). No que diz respeito a gestão dos destinos, os modelos possibilitam monitorar, medir e otimizar a experiência do turista de forma mais eficaz (Ghesh et al., 2024), além de fornecer feedback imediato em relação às reações dos clientes a partir de diferentes abordagens. Dessa forma, contribuem para criar produtos turísticos ou avaliar, com maior celeridade e menor esforço, os resultados de mudanças efetuadas, acompanhando o desempenho da gestão de destinos turísticos, conforme Al-Hyari et al. (2024) e Kim et al. (2024).

Todavia, para que atinja os propósitos almejados, a avaliação e validação da predição para modelos de AM são importantes para medir efetivamente o desempenho dos modelos considerando novos dados, especialmente na automatização da avaliação da experiência turística, principal objetivo em uma predição (Witten et al., 2017). Dessa forma, este estudo adensou o rol de pesquisas que visam identificar variáveis e dimensões para compreender a experiência holística geral dos turistas, lacuna apontada por Kabadayi et al. (2019).

Por meio dos passos apresentados na metodologia deste trabalho e dos testes realizados, foi possível avaliar os modelos que melhor se adaptam para a classificação e predição dos domínios da experiência. O modelo *Bayes* teve maior eficiência na classificação da base unificada Jalapão/Superagui, se comparado com os desempenhos obtidos na classificação do Jalapão e nas predições do Superagui. A eficiência computacional do modelo *Bayes*, aliada ao seu alto desempenho, o torna não apenas o mais preciso, mas também o modelo mais viável e prático para executar predições de novas bases de dados.

A contribuição prática da pesquisa, portanto, consiste em disponibilizar metodologia testada e validada que pode ser replicada em distintos momentos para analisar dados referentes a um determinado atrativo e comparar os indicadores a partir de mudanças nas estratégias implementadas, conforme Ghesh et al., (2024) e Kim et al. (2024).

Todavia, como visto na literatura, o AM depende da constante interação entre o algoritmo e o pesquisador, portanto, ainda há possibilidade de aprimorar os modelos, o que leva à necessidade de novas incursões no tema para aprimorar a metodologia. Nesse sentido, sugere-se replicar a metodologia aqui apresentada para analisar dados de outros destinos naturais, a fim de melhorar o desempenho dos modelos e aprofundar a aprendizagem considerando o maior volume de dados, consoante com a indicação de Witten et al. (2017).

Uma limitação do estudo já citada anteriormente foi a impossibilidade de aplicar as duas fases previstas com todos os modelos. Nesse sentido, sugere-se estudos futuros envolvendo o balanceamento dos dados e a realização de novas avaliações com maior poder computacional, principalmente para os modelos RNA, SVM e Bi-LSTM. Sugere-se ainda, uma avaliação do desempenho dos modelos considerando outra métrica principal, como o F1-score, que pode ser mais representativo e robusto para classes desbalanceadas.

Em síntese, este trabalho vai além da validação técnica de algoritmos, disponibilizando aos gestores do setor turístico, uma ferramenta estratégica capaz de converter a voz do cliente, dispersa em grandes volumes de textos, em subsídios relevantes para a tomada de decisão. O assunto, porém, não se esgota, há ainda um grande percurso a ser percorrido até que as ferramentas de IA possam contribuir de forma efetiva para a solução dos problemas no turismo, objetivo para o qual este estudo se propôs a contribuir.

Cabe ressaltar que as métricas aqui obtidas não seriam aplicáveis em casos complexos como, por exemplo, diagnósticos de doenças e validação da identidade de usuário, pois é necessário considerar o contexto, o

risco de resultados inconsistentes e o impacto do erro. Para uma conclusão aceitável na determinação efetiva da qualidade de um modelo, deve-se ter mais informações e um contexto adicional aos valores de acurácia (Witten et al., 2017).

Referências

- Al-Hyari, H. S., I. M., Al-Smadi & S. R., Weshah (2024). The impact of artificial intelligence (ai) on guest satisfaction in hotel management: an empirical study of luxury hotels. *GeoJournal of Tourism and Geosites*, 48(2), 810–819. <https://doi.org/10.30892/gtg.482spl15-1081>
- Bandeira, L. D. C. L., & de Menezes, P. D. L. (2022). A experiência dos hóspedes na hotelaria de João Pessoa: Uma análise a partir do conteúdo gerado no website Booking.com. *Revista Turismo em Análise*, 33(1), 191–212. <https://doi.org/10.11606/issn.1984-4867.v33i1p191-212>
- Beni, M. C. (2019). *Análise estrutural do turismo*. São Paulo, SP: Senac.
- Caracristi, M. de F. de A., Feger, J. E., Silva, T. M. da, & Marynowski, J. E. (2020). Uma viagem pelo Jalapão, Brasil: Análise das experiências turísticas. *Revista Paranaense de Desenvolvimento (RPD)*, 41(138), 89–110. <https://ipardes.emnuvens.com.br/revistaparanaense/article/view/1164>
- Caracristi, M. de F. de A., Feger, J. E., Marynowski, J. E., & Minasi, S. M. (2021). A demanda turística do Parque Estadual do Jalapão (PEJ, TO, Brasil) baseada em comentários de redes sociais. *Revista Brasileira de Ecoturismo*, 14(3), 291–314. <https://periodicos.unifesp.br/index.php/ecoturismo/article/view/11406/8604>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Claesen, M., & Moor, B. D. (2015). Hyperparameter search in machine learning. *arXiv preprint arXiv:1502.02127*. <https://doi.org/10.48550/arXiv.1502.02127>
- Cooper, C., Fletcher, J., Fyall, A., Gilbert, D., & Wanhill, S. (2007). *Turismo: princípios e práticas (3a ed)*. Porto Alegre: Bookman. https://www.researchgate.net/publication/277230338_Turismo_da_economia_de_servicos_a_economia_da_experiencia
- Correia, L., Lopes, D., Porto, J. V., Lacerda, Y. F., Correia, V. C. A., Bagano, G. O., Pontes, B. S. B., de Melo, M. H. V., Silva, T. E. A., & Meireles, A. C. (2021). Validação de um algoritmo de inteligência artificial para a predição diagnóstica de doença coronariana: Comparação com um modelo estatístico tradicional. *Arquivos Brasileiros de Cardiologia*, 117, 1061–1070. <https://doi.org/10.36660/abc.20200302>
- Crescencio, M., Gonçalves, A. L., & Todesco, J. L. (2020). Processo de classificação de texto: Análise de sentimento das opiniões no Tripadvisor sobre a atração Oktoberfest Blumenau. *Anais do Congresso Internacional de Conhecimento e Inovação–ciki*, 1(1). <https://doi.org/10.48090/ciki.v1i1.867>
- Davenport, T. H. (2018). *The AI advantage: How to put the artificial intelligence revolution to work*. MIT Press.
- De Castro, L. N., & Ferrari, D. G. (2017). *Introdução à mineração de dados*. Saraiva Educação S/A.
- Feger, J. E.; Silva, A. A. N.; Medeiros, M. L. & Brambatti, L. E. (2022) Consumo da gastronomia Caiçara no litoral sul de São Paulo: uma análise a partir dos comentários dos usuários do TripAdvisor. *Revista de Cultura e Turismo*, 16, 01–30. <https://doi.org/10.36113/cultur.v16i2.3408>
- Feger, J. E., Marynowski, J. E., Botta, E. D., & Caracristi, M. de F. de A. (2024a). Experiência vivenciada por turistas no Fervedouro do Ceíça, Tocantins (TO), Brasil. *Revista Paranaense de Desenvolvimento - RPD*, 44(145). <https://ipardes.emnuvens.com.br/revistaparanaense/article/view/1269>
- Feger, J. E., Marynowski, J. E., Reck, S. B., Garcia, R. di F. R., & Caracristi, M. de F. de A. (2024b). Experiência turística no atrativo Serra do Espírito Santo situado no Parque Estadual do Jalapão (TO). *Revista Brasileira de Ecoturismo (RBEcotur)*, 17(1). <https://doi.org/10.34024/rbecotur.2024.v17.14930>

- Ghesh, N., Alexander, M. & Davis, A. (2024). The artificial intelligence-enabled customer experience in tourism: a systematic literature review. *Tourism Review*, 79(5), 1017–1037. <https://doi.org/10.1108/TR-04-2023-0255>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning: Adaptive computation and machine learning series*. MIT Press.
- Guy, I., Mejer, A., Nus, A., & Raiber, F. (2017). Extracting and ranking travel tips from user-generated reviews. *Proceedings of the 26th International Conference on World Wide Web*, 987–996. <https://doi.org/10.1145/3038912.3052632>
- Instituto Chico Mendes de Conservação da Biodiversidade [ICMBio] (2024). Parque Nacional do Superagui. <https://www.gov.br/icmbio/pt-br/assuntos/biodiversidade/unidade-de-conservacao/unidades-de-biomas/marinho/lista-de-ucs/parna-do-superagui/informacoes-sobre-visitacao-2013-parna-do-superagui>.
- Jurman, G., Riccadonna, S., & Furlanello, C. (2012). A Comparison of MCC and CEN error measures in multi-class prediction. *PLoS ONE*, 7(8), Article e41882. <https://doi.org/10.1371/journal.pone.0041882>
- Kaizer, E. F., Caracristi, M. de F. de A., Feger, J. E., Marynowski, J. E., & da Silva, T. M. (2021). Análise da experiência relatada pelos turistas ao visitar o Parque Estadual do Jalapão (PEJ) – TO, Brasil. *Ateliê do Turismo*, 5(1), 183–204. <https://periodicos.ufms.br/index.php/adturismo/article/view/12354>
- Kabadayi, S., Ali, F., Choi, H., Joosten, H. & Lu, C. (2019). Experiência de serviço inteligente em serviços de hospitalidade e turismo, uma conceituação e agenda de pesquisa futura. *Journal of Service Management*, 30(3), 326–48. <https://doi.org/10.1108/josm-11-2018-0377>
- Kim, H. So, K. K. F., Shin, S. & Li, J. (2024). Artificial Intelligence in Hospitality and Tourism: Insights from Industry Practices, Research Literature, and Expert Opinions. *Journal of Hospitality & Tourism Research*, 1–20. <https://doi.org/10.1177/10963480241229235>
- Kozinets, R. V. (2014). *Netnografia: realizando pesquisa etnográfica online*. São Paulo, Penso Editoras
- Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). Foundations of data imbalance and solutions for data democracy. *Data democracy*, 83–106. Academic Press. <https://doi.org/10.1016/B978-0-12-818366-3.00005-8>
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Springer. <https://doi.org/10.1007/978-3-031-02145-9>
- Magalhães, A. M. d., Baracho, R. M. A., & Mandl, T. L. (2021). *Organização da informação: Um modelo semiautomático de classificação de atrações em perfis turísticos usando aprendizado de máquina* [Tese de Doutorado, Universidade Federal de Minas Gerais]. Repositório Institucional da UFMG. <https://repositorio.ufmg.br/handle/1843/35823>
- Manz, C. R., Chen, J., Liu, M., Chivers, C., Regli, S. H., Braun, J., Draugelis, M., Hanson, C. W., Shulman, L. N., Schuchter, L. M., O'Connor, N., Bekelman, J. E., Patel, M. S., & Parikh, R. B. (2020). Validation of a machine learning algorithm to predict 180-day mortality for outpatients with cancer. *JAMA Oncology*, 6(11), 1723–1730. <https://doi.org/10.1001/jamaoncol.2020.4331>
- Marynowski, J. E., Fontana, R. M., Feger, J. E., & Lima, J. D. S. (2023). Domínios de experiência no turismo: Uma análise de comentários turísticos do Jalapão baseada em inteligência artificial. *XX Seminário Anual da Associação Nacional de Pesquisa e Pós-Graduação em Turismo*. Artigo 3382. <https://www.anptur.org.br/anais/anais/sumario.php?versao=20>
- Maschio, J. B., & Marynowski, J. E. (2022). *Processamento de linguagem natural na classificação dos domínios de experiência* [Monografia de Especialização, Universidade Federal do Paraná]. Biblioteca Digital: Trabalhos de Especialização. <https://acervodigital.ufpr.br/xmlui/handle/1884/82178>
- McKenzie, G., & Adams, B. (2018). A datadriven approach to exploring similarities of tourist attractions through online reviews. *Journal of Location Based Services*, 12(2), 94–118. <https://doi.org/10.1080/17489725.2018.1493548>

- Ministério do Turismo. (2024). *Revista eletrônica tendências do turismo*. <https://www.gov.br/turismo/pt-br/aceso-a-informacao/acoes-e-programas/rede-inteligencia-mercado/revista-tendencias-2024-vfinal.pdf>
- Mondo, T. S., Soares, M. H. A., de Lamare Leite, F. C., & Toner, R. (2020). Análise da qualidade de serviços turísticos na Fortaleza de São José da Ponta Grossa – Florianópolis, utilizando o Tourqual. *Revista Turismo em Análise*, 31(3), 477–498. <https://doi.org/10.11606/issn.1984-4867.v31i3p477-498>
- Paes-Cesário, M. F., & Melo, A. S. Q. (2023). Inteligência artificial e a atividade turística: possibilidades de melhorias na experiência ou dificuldades na usabilidade? *XX Seminário Anual da Associação Nacional de Pesquisa e Pós-Graduação em Turismo*. Artigo 3532. <https://www.anptur.org.br/anais/anais/sumario.php?versao=20>
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1–135. <https://doi.org/10.1561/15000000011>
- Pereira, A. S., Shitsuka, D. M., Parreira, F. J., & Shitsuka, R. (2018). *Metodologia da Pesquisa Científica*. Santa Maria, UFSM. <http://repositorio.ufsm.br/handle/1/15824>
- Pine, B. J., & Gilmore, J. H. (1999). *The experience economy: Work is theatre & every business a stage*. Harvard Business School Press. <https://doi.org/10.5860/choice.37-2254>
- Rose, D. (2020). *Fundamentos da inteligência artificial: Aprendizado de máquina*. Editora LinkedIn. <https://books.google.com.br/books?id=Pwz3zwEACAAJ>
- Silva, F. F., Barreto, L. M. T. da S., & Marques Júnior, S. (2021). Dimensões da experiência do cliente em parques temáticos com base nas avaliações de viagens online. *Revista Turismo em Análise*, 32(3), 532–553. <https://doi.org/10.11606/issn.1984-4867.v32i3p532-553>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Souza, S. C. de., & Machado, D. F. C. (2017). Uso e influência das mídias sociais no planejamento de viagens: Um estudo quantitativo. *Revista Turismo em Análise*, 28(2), 254–270. <https://doi.org/10.11606/issn.1984-4867.v28i2p254-270>
- Tocantins. Gestão das Unidades de Conservação do Tocantins. <https://gesto.to.gov.br/uc/45/>
- Tsur, A., Batsry, L., Toussia-Cohen, S., Rosenstein, M. G., Barak, O., Brezinov, Y., Yoeli-Ullman, R., Sivan, E., Sirota, M., Druzin, M. L., Stevenson, D. K., Blumenfeld, Y. J., & Aran, D. (2020). Development and validation of a machine-learning model for prediction of shoulder dystocia. *Ultrasound in Obstetrics Gynecology*, 56, 588–596. <https://doi.org/10.1002/uog.21878>
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). *Data mining: practical machine learning tools and techniques*. Elsevier. <https://doi.org/10.1016/C2015-0-02071-8>
- Yang, L., Hanneke, S., & Carbonell, J. (2013). A theory of transfer learning with applications to active learning. *Machine learning*, 90, 161–189. <https://doi.org/10.1007/s10994-012-5310-y>