Customer Experiential Knowledge Classification in Online Tourist Reviews with a Large Language Model

Alessandro Costa Ribeiro[1,3], Emerson Cleister Lima Muniz[2], Giovani Gracioli[1], João Artur Souza[1]

[1]Universidade Federal de Santa Catarina (UFSC) - Florianópolis / SC - Brasil
[2]Universidade Federal de Sergipe (UFS) - São Cristóvão / SE - Brasil
[3]Universidade do Estado de Mato Grosso (UNEMAT) - Nova Mutum / MT - Brasil

acr.ale@gmail.com, eng.prod.emerson@gmail.com, giovanig@gmail.com, jartur@egc.ufsc.br

Abstract. *Online travel reviews are a primary source of customer knowledge for tourism destinations, yet they are usually analyzed only through sentiment polarity. This paper evaluates the use of a large language model (LLM) to classify reviews in Portuguese according to the six dimensions of Customer Experiential Knowledge (CEK). Reviews from TripAdvisor and Google Reviews about attractions and businesses of Chapada dos Guimarães (Mato Grosso, Brazil), an emerging nature-based destination, were automatically classified and compared with an independent human labeling of 300 reviews. Agreement on the dominant dimension reached Cohen's kappa of 0.670, and the binary salience analysis achieved a macro F1 of 65.7%. The results indicate that LLM-based CEK classification is a viable instrument to profile tourist experiences and to support evidence-based destination management.*

Resumo. *As avaliações de viagem on-line são fonte primária de conhecimento do cliente para destinos turísticos, mas costumam ser analisadas apenas pela polaridade do sentimento. Este artigo avalia o uso de um modelo de linguagem de larga escala (LLM) para classificar avaliações em português segundo as seis dimensões do Conhecimento Experiencial do Cliente (CEK). Avaliações do TripAdvisor e do Google Reviews sobre atrativos e negócios de Chapada dos Guimarães (Mato Grosso), destino brasileiro emergente de base natural, foram classificadas automaticamente e confrontadas com rotulagem humana independente de 300 avaliações. A concordância na dimensão dominante atingiu Kappa de Cohen de 0,670 e a análise binária de saliência alcançou F1 macro de 65,7%. Os resultados indicam que a classificação CEK baseada em LLM é instrumento viável para traçar o perfil experiencial turístico e apoiar a gestão de destinos orientada por evidências.*

1. Introdução

O paradigma do turismo inteligente sustenta-se na coleta, na troca e no processamento intensivo de dados gerados ao longo da experiência turística, convertidos em conhecimento capaz de amparar a tomada de decisão dos atores do destino [Gretzel et al. 2015]. As avaliações publicadas espontaneamente por viajantes em plataformas como TripAdvisor e Google Reviews constituem uma das fontes mais ricas e acessíveis de conhecimento do cliente, sobretudo para destinos emergentes que não dispõem de

infraestrutura própria de pesquisa de demanda [Del Vecchio et al. 2018]. A literatura de Customer Knowledge Management (CKM) aplicada ao turismo reconhece essas manifestações como conhecimento do cliente, isto é, conhecimento produzido pelo turista sobre sua própria experiência, cuja captura e análise sistemáticas alimentam processos de inovação e de gestão inteligente da experiência [Muniz et al. 2021, Khan and Shao 2024].

A prática analítica corrente, contudo, tende a reduzir esse material à polaridade do sentimento, classificando avaliações como positivas, negativas ou neutras. Embora útil para o monitoramento de reputação, a polaridade pouco informa sobre o conteúdo experiencial da avaliação: duas avaliações igualmente positivas podem expressar registros completamente distintos, como o deslumbramento sensorial diante de uma paisagem e a aprovação normativa da organização de um atrativo. Para a gestão do conhecimento do cliente, interessa saber não apenas se o turista aprovou, mas que tipo de experiência ele viveu e relatou. O construto de Conhecimento Experiencial do Cliente (CEK, na sigla em inglês), proposto por Letto et al. (2021), oferece um arcabouço para essa leitura ao decompor o conhecimento tácito expresso nas narrativas de consumo em seis dimensões, três de natureza interna ao indivíduo (instintiva, epicurista e transcendental) e três de natureza externa, socioculturalmente contextualizadas (normativa, subcultural e simbólica).

Operacionalizar o CEK em escala, porém, exige classificar automaticamente milhares de textos livres, curtos, informais e multilíngues, tarefa para a qual abordagens supervisionadas clássicas demandariam grandes conjuntos rotulados inexistentes nesse domínio. Os modelos de linguagem de larga escala (LLMs) tornam viável a classificação multidimensional a partir de definições conceituais fornecidas em linguagem natural, sem treinamento específico, mas seu emprego em tarefas de classificação de conhecimento experiencial em português permanece pouco explorado e carece de validação empírica contra julgamento humano. Revisão integrativa recente de 51 estudos empíricos sobre Big Data Analytics aplicado à experiência turística, publicados entre 2013 e 2025, confirma essa lacuna: o tratamento textual é dominado pela análise de agrupamentos (35,3% dos estudos) e de sentimentos (31,4%) e, entre os 34 estudos publicados a partir de 2022, apenas um recorre a arquitetura Transformer e nenhum incorpora IA generativa; constructos que decompõem o conteúdo experiencial do conhecimento do cliente, como as dimensões do CEK, permanecem sem operacionalização computacional [referência omitida para revisão].

Este artigo avalia a viabilidade do uso de um LLM para classificar avaliações turísticas on-line segundo as seis dimensões do CEK, tendo como campo empírico Chapada dos Guimarães (Mato Grosso), destino turístico emergente brasileiro de base natural. As contribuições são três: (i) a operação do construto CEK como esquema de classificação automática de avaliações, deslocando a análise da polaridade para o conteúdo experiencial; (ii) um protocolo de validação que combina concordância na dimensão dominante, mensurada pelo Kappa de Cohen, e análise binária por dimensão, mensurada por precisão, revocação e F1; e (iii) evidências empíricas sobre os acertos e os padrões de erro do modelo em corpus em português, com implicações práticas para a gestão de destinos. O estudo está organizado da seguinte forma, a Seção 2 apresenta a fundamentação teórica, a Seção 3 descreve o método, a Seção 4 reporta e discute os resultados, a Seção 5 examina as limitações e a Seção 6 conclui e aponta trabalhos futuros.

2. Fundamentação Teórica

2.1. Conhecimento do cliente e Big Data no turismo inteligente

O turismo inteligente é caracterizado pela aplicação intensiva de tecnologias de informação e comunicação à cadeia de valor turística, articulando destinos, negócios e experiências inteligentes em um ecossistema interdependente cuja base comum é o dado [Gretzel et al. 2015, Boes et al. 2016]. Nesse ecossistema, o componente dos negócios inteligentes é apontado como o menos explorado pela literatura [Gretzel et al. 2015, Ballina 2022], embora seja justamente nele que o conhecimento do cliente se converte em decisões operacionais de produtos e serviços. A literatura de CKM no turismo descreve três fluxos de conhecimento, o conhecimento sobre o cliente, o conhecimento para o cliente e o conhecimento do cliente, e situa nas experiências compartilhadas on-line a manifestação mais abundante do terceiro fluxo [Muniz et al. 2021]. Estudos empíricos demonstram a extração de conhecimento do cliente a partir de mídias sociais e plataformas de avaliação mediante a combinação de netnografia e business analytics [Del Vecchio et al. 2020], e análises do determinante tecnológico em destinos inteligentes recorrem à classificação automática de conteúdo avaliativo para inferir qualidade percebida [Díaz-González et al. 2022]. Em comum, esses trabalhos tratam a avaliação on-line como ativo informacional de Big Data, caracterizado por volume, variedade e velocidade, cuja conversão em conhecimento acionável depende de técnicas de Processamento de Linguagem Natural (PLN) [Khan and Shao 2024]. Mapeamento recente da literatura confirma a centralidade das avaliações on-line como fonte: plataformas de avaliação e mídias sociais respondem, respectivamente, por 49,0% e 33,3% dos estudos que analisam a experiência turística por Big Data Analytics, com o TripAdvisor como fonte mais recorrente, embora o tratamento textual desses dados permaneça majoritariamente restrito à polaridade do sentimento [referência omitida para revisão].

2.2. O Conhecimento Experiencial do Cliente e suas dimensões

Letto et al. (2021) definem o CEK como o conhecimento tácito que os clientes têm sobre o modo como vivem suas experiências de consumo, moldado tanto pelo contexto sociocultural em que estão inseridos quanto por impressões subjetivas profundamente entranhadas no relato da experiência. A partir de um estudo netnográfico, os autores decompõem esse conhecimento em seis dimensões, organizadas em dois eixos. O eixo interno reúne aquilo que o turista sente por conta própria diante da experiência. A dimensão instintiva corresponde às reações imediatas e sensoriais, anteriores a qualquer reflexão, como o espanto diante de uma paisagem ou o frescor de uma cachoeira. A dimensão epicurista expressa o prazer de apreciar e desfrutar, manifestado na beleza de um cenário, no sabor de um prato ou no conforto de uma hospedagem. A dimensão transcendental abrange vivências de sentido mais profundo, de contemplação, paz interior ou transformação pessoal, como a sensação de reconexão proporcionada pela natureza. O eixo externo reúne aquilo que o turista interpreta a partir de referências sociais e culturais compartilhadas. A dimensão normativa traduz julgamentos sobre o que se espera de um bom serviço, baseados em normas e padrões de qualidade, como observações sobre limpeza, organização, atendimento ou custo-benefício. A dimensão subcultural manifesta o pertencimento a grupos com gostos e códigos próprios, como trilheiros, surfistas ou observadores de aves, que se reconhecem pela linguagem e pelas práticas que partilham. A dimensão simbólica diz respeito aos significados de identidade e distinção associados à experiência, ou seja, ao que visitar determinado lugar comunica sobre quem é o turista,

seu status ou seu estilo de vida. Para a gestão turística, ler as avaliações por essas dimensões permite identificar o perfil experiencial dominante de cada atrativo ou negócio, uma informação qualitativamente distinta da nota média e da polaridade do sentimento e diretamente acionável no desenho de produtos, na comunicação e na priorização de melhorias.

3. Método

Este estudo adota uma abordagem empírica e quantitativa, de caráter exploratório, voltada a avaliar a viabilidade de um modelo de linguagem de larga escala como instrumento de classificação automática de avaliações turísticas on-line segundo as seis dimensões do Conhecimento Experiencial do Cliente. O delineamento compreende quatro etapas, a constituição de um corpus de avaliações públicas, a classificação automática das dimensões CEK pelo LLM, a construção de um padrão de referência por rotulagem humana independente e a aferição da concordância entre as classificações automática e humana por meio de medidas estatísticas. As subseções seguintes detalham o contexto empírico e o corpus (Seção 3.1), o procedimento de classificação automática (Seção 3.2) e o protocolo de validação (Seção 3.3).

3.1. Contexto empírico e corpus

O campo empírico do estudo é Chapada dos Guimarães - Mato Grosso, um destino turístico emergente brasileiro de base natural, cujos atrativos e negócios concentram intensa produção espontânea de avaliações on-line. O corpus é composto por avaliações públicas extraídas do TripAdvisor e do Google Reviews relativas aos principais atrativos e negócios do destino, entre eles o Parque Nacional da Chapada dos Guimarães, gerido pelo ICMBio, tais como parques e atrativos naturais, restaurantes, meios de hospedagem e agências de receptivo. A coleta combinou consumo de APIs públicas e ferramentas de extração, registrando texto livre, nota numérica, data, idioma e metadados do estabelecimento. Os dados passaram por limpeza (remoção de caracteres de controle e normalização Unicode), detecção e tradução automatizada de avaliações em idiomas estrangeiros e preparação textual com tokenização, lematização e remoção de stopwords contextualizadas ao domínio turístico. Identificadores dos avaliadores foram anonimizados em conformidade com a Lei Geral de Proteção de Dados.

3.2. Classificação automática das dimensões CEK

A classificação foi realizada pelo LLM google/gemini-2.5-flash, instruído com as definições conceituais das seis dimensões de Letto et al. (2021) e exemplos de referência. Para cada avaliação, o modelo atribui seis pontuações independentes e não normalizadas no intervalo [0, 1], uma por dimensão, interpretadas em três faixas: de 0,0 a 0,3, dimensão ausente ou irrelevante; de 0,3 a 0,6, dimensão moderadamente presente; e de 0,6 a 1,0, dimensão central ao discurso do avaliador. A independência das pontuações configura uma classificação multi-rótulo, coerente com a natureza do construto, uma vez que uma mesma narrativa pode mobilizar simultaneamente registros sensoriais, normativos e simbólicos. A classificação é aplicada individualmente a cada avaliação, e as pontuações resultantes são agregadas por atrativo para compor a distribuição média das dimensões em cada estabelecimento.

3.3. Protocolo de validação

Para aferir a confiabilidade da classificação automática, uma amostra de 300 avaliações foi rotulada por um avaliador humano de forma independente, isto é, sem acesso prévio à saída do modelo, constituindo o padrão de referência. Duas análises complementares foram conduzidas. A primeira incide sobre a dimensão dominante, definida como aquela de maior pontuação entre as seis em cada avaliação, verificando a frequência com que humano e modelo convergem na mesma atribuição; a concordância foi mensurada pelo coeficiente Kappa de Cohen [Cohen 1960], que desconta a parcela de coincidência atribuível ao acaso e oferece, por isso, medida mais rigorosa do que a taxa bruta de acerto. A segunda análise adota critério binário de saliência: cada dimensão é tratada como uma questão independente de presença, considerando-se saliente toda pontuação igual ou superior a 0,60, e o desempenho é avaliado por precisão, revocação e F1 por dimensão, além das médias macro não ponderadas, que atribuem peso idêntico a cada dimensão independentemente de sua frequência no corpus.

4. Resultados e Discussão

Na análise da dimensão dominante, humano e modelo convergiram em 76,3% das 300 avaliações, o que corresponde a um Kappa de Cohen de 0,670 após o desconto da concordância fortuita, valor situado na faixa de concordância substancial. A Tabela 1 apresenta o desempenho da análise binária de saliência por dimensão, que resultou em F1 macro de 65,7% e acurácia macro de 91,7%. Para ler a tabela, basta ter em mente o que cada medida expressa: a precisão indica, entre as avaliações em que o modelo apontou uma dimensão, a proporção em que ele acertou; a revocação indica, entre as avaliações em que a dimensão realmente estava presente, a proporção que o modelo conseguiu encontrar; o F1 resume em um único número o equilíbrio entre as duas; e a acurácia considera todos os acertos, inclusive quando o modelo corretamente deixa de atribuir uma dimensão que de fato não estava na avaliação.

Tabela 1. Desempenho da classificação CEK por dimensão

Dimensão CEK	Eixo	F1	Precisão	Revocação	Acurácia
Instintivo	Interno	53,3%	100,0%	36,4%	97,7%
Epicurista	Interno	59,3%	95,4%	43,1%	71,7%
Transcendental	Interno	81,5%	78,6%	84,6%	96,7%
Normativo	Externo	85,0%	84,2%	85,7%	94,3%
Subcultural	Externo	60,4%	45,7%	88,9%	93,0%
Simbólico	Externo	54,5%	46,2%	66,7%	96,7%
Macro (média não ponderada)	–	**65,7%**	–	–	**91,7%**

Lendo a tabela dimensão a dimensão, o desempenho mostrou-se mais equilibrado na normativa (F1 de 85,0%) e na transcendental (F1 de 81,5%), justamente as que reúnem vocabulário mais característico tal como julgamentos de qualidade e organização, no primeiro caso, e expressões de contemplação e significado, no segundo o que facilita o reconhecimento pelo modelo. A dimensão instintiva chama atenção por combinar precisão máxima (100%) com baixa revocação (36,4%) de modo que quando o modelo aponta essa dimensão ele está sempre certo, mas deixa de identificar boa parte das ocorrências reais, sobretudo em avaliações muito curtas, por vezes de uma única palavra, ou que recorrem a gírias e expressões próprias de grupos. O movimento oposto ocorre

nas dimensões subcultural e simbólica, que o modelo tende a atribuir em excesso, daí suas precisões mais baixas (45,7% e 46,2%). Por fim, a acurácia macro de 91,7% parece elevada porque também conta os verdadeiros negativos das muitas avaliações em que a dimensão corretamente não aparece, de modo que ela deve ser lida em conjunto com o Kappa e o F1, indicadores mais sensíveis à efetiva detecção das dimensões presentes.

A Tabela 2 apresenta a matriz de confusão da análise de dimensão dominante, que permite localizar a origem das discordâncias. Nessa matriz, cada linha corresponde à dimensão atribuída pelo avaliador humano e cada coluna à dimensão escolhida pelo modelo, neste sentido os valores na diagonal são as avaliações em que humano e modelo concordam, e os valores fora da diagonal mostram para qual dimensão o modelo desviou quando discordou do humano.

Tabela 2. Matriz de confusão da dimensão dominante (humano × LLM)

Humano \ LLM	Inst.	Epic.	Transc.	Norm.	Subc.	Simb.	Total
Instintivo	**8**	0	0	2	5	0	15
Epicurista	11	**109**	8	9	12	6	155
Transcendental	1	0	**23**	1	0	1	26
Normativo	0	5	0	**61**	4	2	72
Subcultural	0	0	0	1	**21**	0	22
Simbólico	0	1	0	0	1	**8**	10
Total	**20**	**115**	**31**	**74**	**43**	**17**	**300**

As discordâncias concentram-se na dimensão epicurista, o registro mais frequente do corpus (155 das 300 avaliações na rotulagem humana), parte de suas avaliações é distribuída pelo modelo para dimensões vizinhas, sobretudo a instintiva, a subcultural e a transcendental. Esse desvio é compreensível, pois, na experiência turística de base natural, o prazer estético diante das paisagens com frequência se confunde com a reação sensorial imediata e com a contemplação mais profunda, uma fronteira que desafia até avaliadores humanos. As dimensões externas, ao contrário, apresentam diagonais fortes, com 61 das 72 avaliações normativas e 21 das 22 subculturais recuperadas corretamente, o que indica que o modelo as reconhece com segurança.

Do ponto de vista aplicado, os resultados indicam concordância suficiente para sustentar o emprego da classificação CEK como instrumento de leitura do perfil experiencial de atrativos e negócios. Agregadas por estabelecimento, as pontuações evidenciam quais atributos da experiência são mais frequentemente acionados pelos turistas, permitindo, por exemplo, distinguir atrativos de perfil predominantemente contemplativo daqueles avaliados sobretudo por critérios normativos de serviço, e direcionar decisões de comunicação, posicionamento e melhoria orientadas por evidências [Del Vecchio et al. 2020, Khan and Shao 2024].

5. Limitações

As limitações deste estudo estão na interpretação dos resultados tal como o padrão de referência foi produzido por um único avaliador humano; o Kappa reportado mensura, portanto, a concordância entre o modelo e esse avaliador, e não um consenso

interavaliadores, de modo que parte das discordâncias pode refletir a subjetividade intrínseca da tarefa. Segunda, o critério binário de saliência com limiar de 0,60 não captura a proximidade ordinal entre pontuações, penalizando igualmente erros marginais e grosseiros. Terceira, o Kappa é sensível ao desbalanceamento de classes, condição atenuada neste estudo pelo fato de nenhuma dimensão concentrar mais de 80% das ocorrências, mas não eliminada, dado o predomínio da dimensão epicurista. E por fim avaliou-se um único LLM, sem comparação com linhas de base supervisionadas ou com outros modelos, e em um único destino de base natural, o que restringe a generalização dos achados para outros perfis de destino e segmentos turísticos.

6. Conclusão e Trabalhos Futuros

Este artigo avaliou a viabilidade do uso de um modelo de linguagem de larga escala para classificar avaliações turísticas on-line em português segundo as seis dimensões do Conhecimento Experiencial do Cliente, deslocando a análise da polaridade do sentimento para o conteúdo experiencial das narrativas. Validada contra rotulagem humana independente de 300 avaliações, a classificação alcançou concordância substancial na dimensão dominante (Kappa de 0,670) e F1 macro de 65,7% na análise binária de saliência, com padrões de erro interpretáveis e concentrados nas fronteiras semanticamente próximas entre os registros internos da experiência. Os resultados sustentam o emprego da abordagem como instrumento de baixo custo para traçar perfis experienciais de atrativos e negócios em destinos emergentes, contexto pouco contemplado por uma literatura empírica que se concentra em destinos consolidados da Europa e da Ásia [referência omitida para revisão]. A abordagem apoia processos de gestão do conhecimento do cliente baseados em Big Data.

Trabalhos futuros incluem a ampliação do padrão de referência com múltiplos anotadores e mensuração da concordância interavaliadores, comparação do LLM com linhas de base supervisionadas ajustadas para o português e com outros modelos de linguagem, investigação de métricas ordinais e de limiares alternativos de saliência e a replicação do protocolo em destinos de perfis distintos, de modo a verificar a estabilidade das dimensões CEK como esquema de classificação automática da experiência turística.

Referências

Ballina, F. J. (2022) “Smart Business: the element of delay in the future of smart tourism”, Journal of Tourism Futures, v. 8, n. 1, p. 37–54.

Boes, K., Buhalis, D. and Inversini, A. (2016) “Smart tourism destinations: ecosystems for tourism destination competitiveness”, International Journal of Tourism Cities, v. 2, n. 2, p. 108–124.

Cohen, J. (1960) “A coefficient of agreement for nominal scales”, Educational and Psychological Measurement, v. 20, n. 1, p. 37–46.

Del Vecchio, P., Secundo, G. and Passiante, G. (2018) “Analyzing Big Data through the lens of customer knowledge management: evidence from a set of regional tourism experiences”, Kybernetes, v. 47, n. 7, p. 1348–1362.

Del Vecchio, P., Mele, G., Ndou, V. and Secundo, G. (2020) “Detecting customers knowledge from social media big data: toward an integrated methodological

framework based on netnography and business analytics", Journal of Knowledge Management, v. 24, n. 4, p. 799–821.

Díaz-González, S. et al. (2022) "Strategic technological determinant in smart destinations: obtaining an automatic classification of the quality of the destination", Industrial Management and Data Systems, v. 122, n. 10, p. 2299–2330.

Gretzel, U., Sigala, M., Xiang, Z. and Koo, C. (2015) "Smart tourism: foundations and developments", Electronic Markets, v. 25, n. 3, p. 179–188.

Letto, B., Pascucci, F. and Gregori, G. L. (2021) "Defining customer experiential knowledge and its dimensions: a conceptualization starting from a netnographic study of specialty coffee blogs", Journal of Knowledge Management, v. 25, n. 11, p. 149–174.

Khan, M. N. and Shao, Z. (2024) "Impact of Big Data and Knowledge Management on Customer Interactions and Consumption Patterns: Applied Science Research Perspective", Engineering, Technology and Applied Science Research, v. 14, n. 3, p. 14125–14133.

Muniz, E. C. L. et al. (2021) "Customer knowledge management and smart tourism destinations: a framework for the smart management of the tourist experience – SMARTUR", Journal of Knowledge Management, v. 25, n. 5, p. 1336–1361.