

Pedro Abel Pestana Morais

Opinion mining and sentiment analysis:

Algarve's online tourist generated content



UNIVERSIDADE DO ALGARVE

FACULDADE DE ECONOMIA

2021

Pedro Abel Pestana Morais

Opinion mining and sentiment analysis:

Algarve's online tourist generated content

Dissertação de Mestrado em Gestão de Marketing

Trabalho realizado sob a orientação de:

Célia Maria Quitério Ramos

Nelson Manuel da Silva de Matos



UNIVERSIDADE DO ALGARVE

FACULDADE DE ECONOMIA

2021

Extração de opiniões e análise de sentimento:

A imagem do destino turístico no caso do Algarve

Declaração de autoria e copyright

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Pedro Abel Pestana Morais

© **Copyright:** Pedro Abel Pestana Morais

A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

ÍNDICE GERAL

ÍNDICE DE TABELAS	vi
ÍNDICE DE FIGURAS	vii
RESUMO	ix
ABSTRACT	x
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	5
2.1. Imagem do destino turístico	5
2.2. Formação da imagem	6
2.3. Agentes de influência	14
2.4. Imagem do destino online	16
2.5. Processamento de linguagem natural	19
2.6. Análise de sentimentos em UGC (User Generated Content)	20
2.6.1. Implementação da análise de sentimentos	22
2.7. Presente estudo	24
2.8. Modelo conceptual	27
3. METODOLOGIA DA INVESTIGAÇÃO	29
3.1. Aplicação do método	29
3.1.1. Recolha de Dados	29
3.1.2. Pré-processamento	30
3.1.3. Extração de características	33
3.1.4. Aplicação do algoritmo da análise de sentimentos	33
3.1.5. Avaliação dos resultados	35
4. ANÁLISE DE RESULTADOS	36
4.1. Imagem cognitiva	36
4.2. Imagem Afetiva	41

4.3. Imagem conativa	43
4.4. Imagem global (holística).....	45
5. DISCUSSÃO	47
5.1. Implicações práticas para a gestão do setor do Turismo	50
5.2. Contributo para a comunidade científica.....	51
6. CONCLUSÃO	53
6.1. Limitações e recomendações	55
6.2. Trabalho futuro	55
REFERÊNCIAS BIBLIOGRÁFICAS	58

ÍNDICE DE TABELAS

Tabela 1.1. Indicadores provisórios do turismo do Algarve - ano de 2020.	2
Tabela 4.1. Média da valência das categorias e análise de palavras-chave por categoria.....	37
Tabela 4.2. Vinte termos afetivos mais relevantes nas críticas positivas.....	41
Tabela 4.3. Vinte termos afetivos mais relevantes nas críticas negativas.....	42
Tabela 4.4. Cinquenta termos mais frequentes.	44
Tabela 4.5. Termos que mencionam recomendação positiva e negativa.	45

ÍNDICE DE FIGURAS

Figura 2.1. Dimensões da Imagem de um Destino turístico	8
Figura 2.2. Inter-relação entre imagem projetada e imagem percebida.....	18
Figura 2.3. Modelo conceptual da formação da imagem do destino turístico	28
Figura 3.1. Estrutura da metodologia para o processamento da linguagem natural .	29
Figura 4.1. Duzentos termos mais frequentes numa <i>Word Cloud</i>	43
Figura 4.2. Imagem holística do Algarve.....	46

Agradecimentos

Deixo o meu profundo agradecimento à minha família, por todo o apoio, motivação, disponibilidade e conselhos. Certamente, tornaram esta jornada mais leve.

Um especial agradecimento à minha irmã, Inês Salomé Morais.

À minha cara-metade, Liliana Salas Crispim, por todo o apoio, companhia e motivação.

Aos meus amigos de infância:

Sérgio Paulo, Pedro Paulo, André Lourenço Fernandes, Jaime Jacob, João Ferreira e Hugo Cavaco, - Pela irmandade e companheirismo, que tornaram os momentos mais amargos em peripécias para contar depois.

Agradeço, especialmente, aos meus orientadores de dissertação de mestrado, pela supervisão, mentoria e cooperação:

Célia Maria Quitério Ramos.

Nelson Manuel da Silva de Matos.

RESUMO

A imagem do destino turístico é um tema que, desde os anos 70, começou a levantar o interesse da comunidade científica. Este conceito tem sido extensivamente estudado, sendo um dos focos de estudo a compreensão do processo de formação da imagem do destino turísticos. Apesar das divergências na literatura quanto às componentes da imagem, existe um vasto consenso quanto ao argumento de que a imagem é um constructo tridimensional, em que as dimensões formam uma estrutura hierárquica com uma relação de influência mútua. O meio digital alterou definitivamente este processo criativo, a imagem passou a ser cocriada pelos turistas, devido à facilidade e volume com que se realiza atividades como o *word-of-mouth*, a partilha de fotografias e vídeos, tornando a imagem percebida num fator fulcral, e retirando a imagem projetada do pedestal da formação da imagem do destino. Assim, novas estratégias e técnicas de análise para explorar o conceito da imagem do destino no meio digital surgiram nas últimas duas décadas face ao surgimento das redes sociais, apresentando um desafio ao tratamento e conhecimento da informação que daqui é proveniente. O presente estudo foca-se na compreensão da imagem do destino no caso do Algarve a partir de críticas de turistas internacionais presentes na Internet, e aplica o conhecimento teórico do conceito de imagem do destino turístico em conjunto com técnicas de processamento de linguagem natural, integrando uma análise de sentimento com conteúdo gerado pelos utilizadores. O sentimento expresso pelos turistas é, maioritariamente, positivo em relação ao Algarve. Os fatores positivos superam os negativos, no entanto, encontram-se situações em que a qualidade do turismo no algarve pode ser melhorada. Esta dissertação contribui para o corpo literário do tema e levanta comentários de interesse para o tecido empresarial do mercado turístico.

Palavras-chave: Imagem do destino turístico; Comportamento do consumidor; Redes sociais; Processamento de linguagem natural; Análise de sentimento; Turismo no Algarve.

ABSTRACT

The tourism destination image is a topic that began to raise the interest of the scientific community in the seventies. This concept has been extensively studied, being one of the focuses of study the understanding of the tourist destination image's formation process. Despite the divergences in the literature as to the components of the image, there is a broad consensus on the argument that the image is formed according to a three-dimensional construct, in which the dimensions form a hierarchical structure and share a relationship of mutual influence. The digital environment has changed this creative process, the image started to be co-created by tourists, due to the ease and volume with which activities such as word-of-mouth, photo and video sharing are carried out, turning the perceived image into a key factor, and removing the projected image from the pedestal of the formation of the destination image. Thus, new strategies and analytics were applied to explore the concept of destination image in the digital environment. These techniques have emerged in the last two decades due to the insurgence of social media, presenting a challenge to the treatment and knowledge of the information that arose. The present study focuses on the understanding of the destination image in the case of Algarve, from the international tourists' reviews present on the Internet, and applies the theoretical knowledge of the concept of tourist destination image, combined with natural language processing techniques, including the integration of sentiment analysis on user-generated content. The sentiment expressed by tourists is mostly positive towards the Algarve. The positive factors outweigh the negative, however, there are situations in which the quality of Algarve's tourism can be improved. This dissertation contributes to the body of literature on the subject and raises comments of interest to the enterprise's fabric of the tourism market.

Keywords: Destination image; Consumer behaviour; social media; natural language processing; Sentiment analysis; Algarve's tourism.

1. INTRODUÇÃO

1.1. Enquadramento geral

A região do Algarve, situada a sul de Portugal continental, constitui um destino turístico que gera uma grande afluência de estrangeiros e hóspedes nas suas infraestruturas. Devido, não só, à sua diversidade paisagística e ao seu clima, mas também ao desenvolvimento económico, social, territorial e demográfico que aconteceu ao longo dos anos. O algarve é, hoje em dia, um destino altamente competitivo.

O tecido do mercado turístico começou a desenvolver-se a partir dos anos 60 no Algarve. Entre as décadas de 70 e 80, o algarve registou um aumento significativo em dormidas e oferta hoteleira. Em Hunt (1975) é explorado a opinião do consumidor relativamente á experiência vivida em quatro estados diferentes dos Estados Unidos, partindo da premissa de que a perceção que determinado indivíduo tem de um destino influencia a viabilidade do destino enquanto uma região turística. Introduzido, assim, um estudo que procura perceber a influência da imagem percebida na escolha do destino pelo consumidor, com base na localização geográfica e os fatores socioeconómicos do publico, tendo constatado que o processo de decisão pelos consumidores é a base para que a imagem de destino se construa, influenciando fortemente o desenvolvimento turístico dos diferentes locais.

Destaca-se, assim, a importância de explorar a imagem do destino, especialmente em regiões em que o turismo representa a indústria principal para gerar receitas, como o Algarve. Em 2020, em contexto de situação pandémica devido ao Covid-19, o Algarve apresentou uma quebra significativa de fluxo turístico e, conseqüentemente, de receita. A tabela 1.1 demonstra o impacto sofrido pelo Algarve neste contexto:

Tabela 1.1. Indicadores provisórios do turismo do Algarve - ano de 2020.
Adaptado de Região de Turismo do Algarve (2020:2)

Indicadores		2020	Var. 2020/19 (%)
Movimento no aeroporto	Estrangeiros	2,042,907	-76%
	Portugueses	163,656	-64%
	Total	2,206,563	-76%
Dormidas	Total	7,926,068	-62%
Hóspedes	Total	1,114,068	-61%
Proveitos	Totais	465,087,903 €	62%
	Aposento	352,515,003 €	62%

1.2. Problemática de investigação

A *Web 2.0*, com o aparecimento das redes sociais, revolucionou a forma como as pessoas comunicam umas com as outras. Os *mass media* começaram a estar acessíveis a qualquer pessoa com acesso à Internet (Jiang et al., 2021; Ren e Hong, 2017), alterando definitivamente o processo de tomada de decisão dos consumidores (Goldsmith, 2006; Litvin et al., 2008; Mak, 2017; Ye et al., 2011). Para a investigação, o aparecimento destes medias colocou um grande desafio à comunidade científica, - a grande dimensão e volume dos dados (B. Liu, 2010). Para superar esta desafio surgiram novas técnicas de explorar e compreender a imagem do destino, que ao longo dos anos com o desenvolvimento da tecnologia foram tornando-se cada vez mais imprescindíveis (Govers e Go, 2008).

De modo a superar o obstáculo apresentado pela dimensão e volume dos dados, é possível aplicar técnicas que levam a inteligência artificial a interagir com a linguagem natural (*Natural Language Processing*). A problemática que se apresenta a esta dissertação é o como aplicar estas técnicas de processamento de linguagem textual, à fonte de dados exorbitante que existe nas redes sociais, de modo a conseguir obter resultados claros e sucintos acerca do constructo complexo que é a imagem do destino (Dwivedi, 2009; Jiang et al., 2021; Marine-Roig e Huertas, 2020; Ren e Hong, 2017; Yu e Zhang, 2020).

De modo a resolver esta problemática, o presente estudo propõe-se a atingir os seguintes objetivos:

- I. Enquadramento da imagem do destino no caso do Algarve, segundo o que está estipulado na teoria presente na literatura acerca da imagem do destino turístico.

- II. Utilizar comentários extraídos do *TripAdvisor* como informação que permite descrever a imagem do destino no caso do Algarve.
- III. Aplicar o processamento de linguagem natural assistido por inteligência artificial. Nomeadamente, fazer uso de pré-processamento textual para extrair características, e algoritmos de análise de sentimento para obter a valência sentimental subjetiva nos comentários, enquadradas na área da análise textual.

1.3. Estrutura da dissertação e Metodologia aplicada

O capítulo 2 apresenta a revisão bibliográfica da literatura utilizada para estudar profundamente o tema da imagem do destino, processamento de linguagem natural e análise de sentimentos. Este capítulo compreende artigos científicos desde a década de 70 até ao momento.

Em primeiro lugar, é explorada a origem e desenvolvimento do conceito de imagem do destino, onde se revê os diferentes contributos que houve no estudo deste fenómeno. Assim, a primeira parte deste capítulo foca-se na definição do conceito, no processo de formação da imagem do destino turístico, em que se refere a possibilidade do conceito se tratar de um constructo composto por três dimensões (cognitiva-afetiva-conativa), e cada uma destas dimensões é debatida consoante o que se descobre na literatura disponível. Por fim, é abordado o conceito da imagem do destino no meio digital, e de que maneira se distingue da imagem do destino tradicional.

Em segundo lugar, introduz-se o tema do processamento de linguagem natural, em que se dá a conhecer as técnicas e conceitos ligados ao método, incluindo a análise de sentimentos. Durante esta secção introduz-se o *text mining*, ou análise textual, onde se faz a ponte entre o método do estudo e o conceito da imagem do destino, abordando o como e porquê de aproveitar o conteúdo criado por utilizadores (UGC – *user generated content*) para conduzir um estudo sobre a imagem do destino turístico. A análise textual permite extrair informação relevante e fazer sentido do conteúdo textual, que, por ser extenso em volume, inclui muito conteúdo vaga e insignificante. Assim, a partir do processamento de linguagem natural, torna-se possível transformar o conteúdo criado pelos utilizadores em categorias e/ou padrões que permite compreender a imagem do destino percebida pelos turistas internacionais.

No capítulo 3 demonstra-se como é que a metodologia considerada no presente estudo, adaptada a partir de outras selecionadas de estudos prévios podem ser aplicadas para atingir os objetivos estipulados para o presente estudo. Descrevem-se as etapas pelas quais os dados extraídos do *TripAdvisor* necessitaram de ser sujeitos para que a informação se tornasse clara e objetiva.

No capítulo 4 apresentam-se os resultados e a respetiva explicação, e de seguida, no capítulo 5 estes dados são discutidos e evidenciam-se as implicações que os resultados obtidos têm no tecido empresarial do mercado turístico, e é refletido qual o contributo que os resultados representam para a comunidade científica.

O capítulo 6 conclui o presente estudo com a reflexão, síntese e comentários finais acerca do mesmo. Terminando, por fim, com as limitações encontradas durante todo o processo da pesquisa e com a proposta de ideias para estudos futuros que possam vir a complementar o que se descobriu acerca da imagem do Algarve enquanto destino turístico.

2. REVISÃO DE LITERATURA

2.1. Imagem do destino turístico

Durante a década de setenta, o conceito de imagem do destino começou a ganhar interesse na comunidade científica. Nomeadamente, Hunt (1975) foi pioneiro no estudo da influência da imagem na escolha do destino pelo consumidor, tendo em conta a localização geográfica e os fatores socioeconómicos do público. Concluiu que a imagem percebida, ao estar intimamente ligada ao processo de decisão pelos consumidores, exerce uma forte influência no desenvolvimento turístico dos locais. Mostrando, assim, que estudos sobre a imagem dos destinos são fulcrais no posicionamento da comunicação de informação turística.

Numa outra perspetiva, Mayo (1975), num estudo focado em parques nacionais como destino turístico, concluiu que fatores como o local de residência, sexo, idade e ocupação não estão significativamente interligados à atratividade aos parques. O que distingue esta investigação da supracitada é o facto que em Mayo (1975) é tido em conta a personalidade, atitude e motivações. Ou seja, fatores psicossociais, assentam na premissa que estes proporcionam uma maior compreensão das razões iniciais que conduzem os indivíduos a querer viajar.

Posteriormente, Crompton (1979) contribuiu para o debate sobre o tema, argumentando que os fatores psicossociais servem meramente para entender o que leva um indivíduo a querer viajar, mas também o que o leva a escolher um destino em específico. O autor refere que, neste aspeto, um destino “exerce” uma força de atração enquanto ponto cultural.

Por sua vez, um destino pode ter uma imagem primária, baseada na realidade, e uma imagem secundária, baseada na perceção/noção que se tem de um determinado destino turístico, que resulta da exposição a informação comunicada por terceiros (Phelps, 1986). Ao dividir o processo de formação da imagem em dois níveis de significado diferentes, demonstra-se que a imagem do destino resulta de duas fontes distintas, em que uma das fontes (secundária) é substituída pela outra (primária) no fim da jornada do turista. Por isso, em Phelps (1986), demonstra-se que a imagem secundária (imagem percebida) tem

de estar em harmonia com a experiência turística que, de facto, se usufrui no destino. Visto que, no caso da formação de uma imagem secundária irreal, o consumidor cria uma expectativa que não se cumpre, tendo um efeito negativo na sua satisfação final. Ou seja, neste estudo (Phelps, 1986) é demonstrado que a imagem do destino se altera consoante a experiência passada no local. A imagem é, por isso, a junção entre elementos induzidos e elementos orgânicos (Gartner e Hunt, 1987).

2.2. Formação da imagem

São vários os estudos que procuram compreender o processo de formação de imagem (Bigné Alcañiz et al., 2009; Dann, 1996; Echtner e Ritchie, 1991; Fakeye e Crompton, 1991; Gartner, 1994; Phelps, 1986), que, no entanto, são consensuais relativamente ao processo de formação. Isto é, a literatura está em conformidade com o que é preconizado por Gunn (1988), quanto às sete fases da experiência de turismo: (1) aglomerado de imagens mentais que auferem a experiências passadas pelo consumidor, que podem ou não estar relacionadas ao(s) destino(s) em seleção; (2) a busca por mais informação para auxiliar a escolha irá alterar essas imagens mentais; (3) decisão final pelo turista sobre o destino de viagem; (4) viagem para o destino; (5) participação ativa nas atividades fornecidas pelo destino; (6) viagem de retorno; e (7) contraste entre a expectativa e a realidade (experiência), o que modifica a imagem que o turista tem do destino. A imagem do destino é, portanto, o resultado deste processo, que está intimamente relacionado com a experiência do turista (Lalicic et al., 2021), e constitui um fator relevante para o sucesso dos locais turísticos (Hunt, 1975; Park et al., 2020). Uma imagem do destino favorável não contribui apenas para o desenvolvimento turístico, mas é também fulcral para os produtos e serviços turísticos disponíveis no local (Nowacki e Niezgoda, 2020).

Na perspetiva de Phelps (1986), a imagem é formada inicialmente a partir de fontes de informação comunicada por terceiros (guias de viagem, panfletos, agentes de viagem, etc.), ou a partir da opinião de outros turistas que tenham conhecimento ou experiência sobre os destinos - “imagem secundária”. No entanto, ao finalizar a viagem, o turista substitui os fatores que conduziram à sua escolha por outros que considerou mais relevantes durante a estadia. Formando, assim, a “imagem primária”.

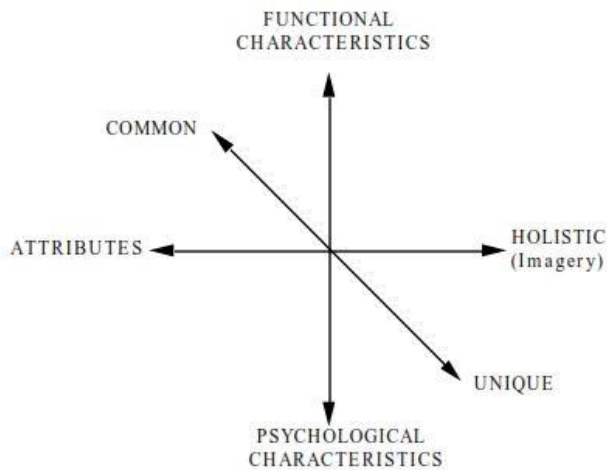
O mesmo é sugerido em Fakeye e Crompton (1991), onde os esforços publicitários conduzem o indivíduo a formar uma imagem orgânica, despoletando a vontade de viajar e a atividade de procurar mais informação sobre o destino e, a partir deste momento, forma-se uma imagem induzida, que depende de fatores psicossociais como a personalidade, experiência, estado de espírito, motivações e necessidades do indivíduo (Crompton, 1979; Mayo, 1975). Por fim, ao viver a experiência turística, a imagem torna-se em algo mais complexo (Fakeye e Crompton, 1991; Phelps, 1986).

Aquando de uma futura viagem, este turista poderá utilizar esta experiência como contraste com a imagem induzida desse novo destino. Assim, em Fakeye e Crompton (1991) propõe-se que a imagem do destino assume três etapas diferentes, (1) imagem orgânica, (2) induzida e (3) complexa.

Dann (1996) argumenta que durante o processo de escolha de destino, a comparação feita entre as opções de viagem apoia-se num processo de *déjà vu*, onde o indivíduo reflete sobre outras experiências reais e de referência. Assim como Fakeye e Crompton (1991), encontram-se argumentos a favor da forte influência de experiências turísticas passadas no processo de formação de imagem do destino.

Echtner e Ritchie (1991) afirmam, quanto ao processo de formação de imagem, que os indivíduos possuem uma imagem de determinado destino, mesmo que não tenham viajado para o local e que não tenham sido expostos a informação publicitária/comercial. Além disso, tendo em conta que a imagem é diferente no antes e no depois da visita, é necessário separar a imagem dos experientes, da imagem dos que nunca experienciaram o destino. Nesta pesquisa (Echtner e Ritchie, 1991) é proposto que a imagem seja composta por três *continuum* dimensionais diferentes, tal como apresentado na Figura 2.1:

Figura 2.1. Dimensões da Imagem de um Destino Turístico



Fonte: Echtner e Ritchie (1991: 43)

- (I) Os atributos dos destinos variam entre comuns e únicos.
- (II) A imagem pode basear-se em atributos que são mensuráveis (*e.g.*, preços e clima) e, como são comuns a todos os destinos, podem ser comparados. Contrariamente, a imagem pode basear-se em sensações holísticas (*e.g.*, sentimentos específicos despoletados por um cenário), algo que destaca um destino em relação a outro.
- (III) As características de um destino podem ser funcionais e, conseqüentemente, tangíveis. Por outro lado, podem ser características psicológicas, que auferem às emoções e ao julgamento subjetivo do turista e, por isso, são intangíveis e imensuráveis.

Vários autores apoiam que as dimensões da imagem do destino dependem e exercem influência entre si, formando um sistema complexo de causa-efeito (Baloglu e Mcclary, 1999; Beerli e Martín, 2004; Carvalho et al., 2020; Gallarza et al., 2002; Gedikoglu et al., 2020).

Em Gartner (1994), são reconhecidos três componentes *hierarquicamente interrelacionados*: a componente cognitiva, a componente afetiva e a componente conativa. Explica, ainda, que estes elementos fazem parte do processo de seleção do destino de viagem. Durante este processo, o consumidor forma uma imagem do destino a partir do que conhece, do que sente e do que o motiva em relação ao local, dependendo

desta avaliação é selecionada uma oportunidade de viagem. Cada componente diz respeito a uma dimensão de significado atribuído ao destino, isto é, a componente cognitiva afere à percepção da realidade, um processo em que o turista contempla os atributos funcionais do destino (Echtner e Ritchie, 1991; Hernández-Mogollón et al., 2014) segundo os seus desejos e motivações, a componente afetiva diz respeito a uma relação de valor que o consumidor atribui ao local, ou seja, se o destino é capaz de oferecer uma experiência que realize os seus desejos (benefícios da compra) (Cardoso et al., 2019), e pode, ou não, entrar em conflito com os valores do consumidor. Por fim, na hierarquia, está a componente da atitude (*action component*), que ocupa o último lugar da cadeia por ser fundamentalmente influenciada pela avaliação cognitiva e afetiva, e envolve a atitude final do consumidor ao decidir um destino de viagem, e a atitude posterior face ao destino (Agapito et al., 2013). Não obstante, a influência da imagem do destino no comportamento do consumidor não se restringe à atitude à priori e à posteriori da experiência, ao qual Agapito et al. (2013) complementa, argumentando que a imagem do destino tem uma forte influência no comportamento do turista durante todo o processo da viagem, desde o momento da tomada da decisão, durante a experiência turística, e durante o após da viagem. Ou seja, a imagem do destino também influencia a possibilidade de haver recomendação pelo destino experienciado, assim como o retorno ao mesmo e, subsequentemente, a lealdade enquanto cliente.

Em Keller (2013) é discutido as associações que os consumidores fazem das marcas. Esta aproximação assemelha-se ao modelo de formação de imagem do destino de Gartner (1994). Estas associações distinguem-se, também, em três partes:

- Atributos – características palpáveis de um determinado serviço ou produto. No caso dos destinos turísticos, estes atributos podem ser as características físicas do mesmo (*e.g.*, clima ou cenário) (Agapito et al., 2013);
- Benefícios – valor pessoal e simbólico atribuído às características. Esta avaliação contempla aquilo que, na mente do turista, o destino lhe pode concretizar, passando pelas características funcionais e não funcionais (Cardoso et al., 2019).
- Atitudes – A avaliação final e global que o turista faz da sua experiência no destino irá condicionar o comportamento posterior à mesma (Cai, 2002). Por sua vez, a imagem do destino formada exerce uma forte influência na lealdade do turista perante o destino (Bigné et al., 2001).

Dum ponto de vista tradicional, a dimensão cognitiva da imagem do destino engloba o conhecimento e crenças acerca de determinado destino (Carvalho et al., 2020), e esta sabedoria inclui uma representação/percepção objetiva e realística dos atributos dos destinos (Baloglu e McCleary, 1999). Como é algo que é relativo ao indivíduo, torna-se numa interpretação subjetiva da realidade, o que significa que difere consoante o turista (Bigné et al., 2001). Por outras palavras, esta dimensão cognitiva da imagem surge durante um processo de avaliação de atributos funcionais do destino, formando-se uma atitude face a um objeto (o destino) a partir da recolha e tratamento de informação. Este processo ocorre segundo uma interpretação e esquematização dos atributos presentes no(s) destino(s), formando uma percepção/crença em torno do local, que pode, posteriormente, ser avaliada segundo as emoções e sentimentos despoletados por esta imagem percebida (Agapito et al., 2013; Hernández-Lobato et al., 2006). Por sua vez, estas emoções poderão estar ligadas a motivações do turista face ao destino.

De acordo com Hernández-Mogollón et al. (2014), estas motivações podem ser de ordem física, como o entretenimento, bem-estar e relaxamento; de ordem cultural, relacionado com a vontade de conhecer melhor outras culturas; de ordem interpessoal, levando o turista a visitar locais para estabelecer novos laços ou a reencontrar-se com amigos e/ou família; e, por último, de ordem social, tornando o destino numa experiência que traz prestígio social, fortificando a sua identidade perante outros indivíduos. Ou seja, atributos como o clima, infraestruturas, história e cultura do local, restaurantes e museus, qualidade do atendimento ou a paisagem, e entre outros, são elementos que influenciam diretamente a avaliação cognitiva, e que contribuem para a imagem do destino (Carvalho et al., 2020; Echtner e Ritchie, 1991; San Martín e Rodríguez del Bosque, 2008). Resumidamente, a imagem na mente do consumidor (turista) é uma representação criada a partir do processamento de estímulos, e esta representação é fortemente influenciada pelos traços psicológicos do turista (Crompton, 1979; Mayo, 1975; San Martín e Rodríguez del Bosque, 2008).

As motivações são o fator inicial que leva o indivíduo a procurar uma experiência turística (Bigné et al., 2001; Fakeye e Crompton, 1991; Gartner, 1994; Mayo, 1975). De modo a compreender o comportamento do consumidor de produtos turísticos, San Martín e Rodríguez del Bosque (2008) refere os fatores *push-pull*. Estes fatores podem ser estímulos internos (*push*) relacionados com as características emocionais/sentimentais do

indivíduo, que expressam os desejos que o atraem para a viagem. Por outro lado, estes fatores podem ser externos (*pull*), que, no fundo, são os atributos do destino que levam à preferência de uma determinada viagem entre o conjunto de possibilidades. Por outras palavras, em Baloglu e McCleary (1999) são definidos dois tipos de fatores diferentes que atuam na formação da imagem do destino: (1) fatores estimulantes (*stimulus factors*) e (2) fatores pessoais (*personal factors*), sendo que o primeiro refere-se ao destino enquanto um objeto palpável e usufruível, o que pode também incluir experiências anteriores (Fakeye e Crompton, 1991); e o segundo está relacionado com as características sociais e psicológicas do consumidor (Mayo, 1975).

Com efeito, a imagem do destino é construída na mente do turista a partir de pistas que lhe são dadas por agentes de formação da imagem, e são selecionadas pelo mesmo (Gartner, 1994; Tasci e Gartner, 2007), o que o leva a balançar entre as características e atributos funcionais de vários destinos em simultâneo (Agapito et al., 2013; Baloglu e McCleary, 1999). Pressupõe-se que a motivação, perante estas características funcionais, está condicionada pelos benefícios atingíveis a partir do destino, e o seu valor pessoal é julgado a partir do estado final desejado pelo turista, o que favorece ou desfavorece a imagem (San Martín e Rodríguez del Bosque, 2008). Trata-se, assim, de um julgamento afetivo (Baloglu e Love, 2005).

A componente avaliativa, que também é conhecida como dimensão afetiva, assim como a dimensão cognitiva, é uma resposta a estímulos que originam das características físicas do destino. Isto é, a componente avaliativa surge a partir de uma avaliação afetiva dos atributos presentes na imagem percebida do destino, ou seja, na dimensão cognitiva da imagem (Agapito et al., 2013; Hernández-Lobato et al., 2006).

Assim como Gartner (1994) indica, a dimensão afetiva da imagem do destino participa numa relação hierárquica interrelacionada. A relação entre as dimensões cognitiva e afetiva exprime-se a partir da seleção de características/atributos relevantes e respetiva avaliação do mesmo segundo as emoções e desejos do turista enquanto indivíduo, tornando-se numa relação de causa-efeito (Baloglu e McCleary, 1999; Beerli e Martín, 2004; Carvalho et al., 2020).

Em Stylidis et al. (2017) é conduzido um estudo pertinente a partir de uma amostra de turistas e residentes de uma cidade na cidade de Eilat no Iraque, que permitiu aplicar o modelo de Gartner (1994) e compreender se, de facto, é possível auferir quanto à influência entre as dimensões da imagem do destino. Neste estudo, foi concluído que, de facto, como supracitado, (1) a dimensão cognitiva está relacionada à imagem afetiva, (2) ambas estas dimensões exercem influência na imagem global, e (3) ambas as dimensões da imagem, juntamente com a imagem global percebida, estão relacionadas com a intenção de recomendação, o que também demonstra que a imagem do destino é um constructo em que o seu significado é algo mais grandioso que a simples junção das suas partes (Agapito et al., 2013). Não obstante, a imagem global (*overall image*) é a que exerce maior influência na lealdade do turista ao destino (Bigné et al., 2001). Estas conclusões estão de acordo com as associações que o consumidor atribui ao produto (Keller, 2013), e colaboram com as conclusões obtidas em Beerli e Martín (2004), que explica que a motivação (imagem cognitiva) para determinada viagem exerce influência na dimensão afetiva da imagem, e que a experiência turística no destino influencia a imagem cognitiva e afetiva do destino. Com efeito, os resultados destas influencias podem traduzir-se na imagem primária e secundária de Phelps (1986), e entra em contacto com aquilo que Fakeye e Crompton (1991) comenta acerca da mutabilidade da imagem do destino, em que a imagem do destino distingue-se em três etapas diferentes na sua formação (imagem induzida – orgânica – complexa). Assume-se, então, que para melhor compreender a imagem do destino, é necessário incluir todas as dimensões da imagem, ao contrário do que é argumentado em Chen (2001), que afirma que os atributos cognitivos são mais práticos e concretos para compreender aquilo que torna um destino único, e que a avaliação afetiva dos atributos da imagem são vagos.

A dimensão conativa da imagem é a manifestação da atitude do consumidor, e pode traduzir-se no comportamento que este tem perante o destino turístico. Isto é, a decisão de embarcar na viagem, a sua atitude consumista no local, a sua lealdade e intenção de recomendação (Woosnam et al., 2020; Zhang et al., 2014). Esta componente da imagem é sucessora e resultado da relação complexa entre a imagem cognitiva e a imagem afetiva do destino (Gartner, 1994).

Há uma ligação evidente entre esta dimensão e as supracitadas, tendo em conta que o comportamento do consumidor está condicionado à realização daquilo que foi levantado

durante o processo de percepção e avaliação do destino (Woosnam et al., 2020). Segundo Stylos et al. (2016), a dimensão conativa compreende o como e o porquê de o conhecimento e sentimentos associados a determinado destino levam o turista a elegê-lo entre todas as outras opções, o indivíduo projeta-se no futuro, contemplando a experiência que o mesmo pretende que se desenvolva.

De acordo com as assunções teóricas do modelo tridimensional da formação da imagem (cognitivo – afetivo – conativo), existirá, portanto, fatores que, independentemente das características e benefícios atingíveis no local, retraem o consumidor de escolher determinado destino e que não estão diretamente relacionados. Um e Crompton (1992) referem que poderão existir atributos restritivos (*constraining attributes*) que são levantados durante o processo de escolha de um destino. Dependendo do número e da avaliação dessas restrições, um destino turístico é elegido para a viagem (Um e Crompton, 1992).

Em Kim et al. (2021) é conduzido um estudo que pretende avaliar se as restrições de viagem (*travel constraints*) condicionam o processo de formação da imagem do destino. Neste estudo (Kim et al., 2021) são considerados três naturezas distintas de restrições (intrapessoal, interpessoal e estrutural), e ao contrastar estes fatores com as diferentes etapas do processo de formação de imagem, o mesmo conclui que quando determinado destino implica restrições a nível intrapessoal irá afetar negativamente a avaliação cognitiva e afetiva da imagem, o que, segundo os vários autores (Agapito et al., 2013; Baloglu e McCleary, 1999; Beerli e Martín, 2004; Bigné et al., 2001; Gartner, 1994; Stylos et al., 2016; Zhang et al., 2014), coloca limitações na escolha do destino, e consequentemente, afetar negativamente a imagem conativa do destino.

Com efeito, a dimensão conativa é também um fator fulcral na atitude futura do turista relativamente ao destino que acaba de experienciar (Afshardoost e Eshaghi, 2020; Stylidis et al., 2017), o que terá impacto na recomendação, lealdade e possibilidade de revisita (Bigné et al., 2001).

Em Papadimitriou et al. (2018) é testado o efeito que as três dimensões têm na propagação de *word-of-mouth* positivo (*i.e* recomendação) a partir de uma amostra de três

grupos distintos de turistas (residentes, turistas com experiência passada no destino e possíveis turistas). Neste estudo (Papadimitriou et al., 2018) é concluído o seguinte:

- A imagem global percebida pelos residentes é maioritariamente influenciada pela dimensão cognitiva da imagem, ao contrário daquilo que é sugerido por Stylidis et al. (2017).
- Possíveis turistas são mais facilmente influenciados pela imagem global do destino, e, respetivamente, pelo *word-of-mouth*.
- A intenção de recomendação pela parte de residentes e turistas passados é influenciada pelos componentes cognitivos e afetivos do destino.
- A dimensão afetiva influencia significativamente a imagem global do destino, o que demonstra a importância do julgamento emocional e da percepção global perante o destino.

Neste aspeto, conseguir que determinado destino realize todas as expectativas do turista (Hosany e Witham, 2010) e, por conseguinte, uma imagem conativa positiva, é possível influenciar um comportamento favorável pela parte do turista e medir, avaliar e compreender o desempenho do local turístico (Afshardoost e Eshaghi, 2020).

2.3. Agentes de influência

Phelps (1986) considera a existência de uma imagem secundária e de outra primária. Em Gartner (1994) é possível encontrar semelhanças, isto é, assentando no argumento de que o processo de formação de imagem é um *continuum* composto por diferentes agentes de ação, é proposta uma definição para cada um desses agentes, que por sua vez se dividem em dois grandes grupos, os agentes induzidos (imagem secundária) e os agentes orgânicos (imagem primária). Assim, é proposta uma etiquetagem para estes agentes, que exercem influência na imagem percebida pelo turista (Gartner, 1994). Estes agentes são denominados e definidos como:

- *Overt Induced I* – Todo o tipo de promoção tradicional, isto é, anúncios de televisão, rádio ou conteúdos impressos, de modo a transmitir aos prospectos uma imagem desejada pelas entidades turísticas dum local específico. Naturalmente, estas entidades têm um interesse em despoletar desejo no recetor.

- *Overt Induced II* – Consiste na informação obtida, seja procurada ou induzida, a partir de entidades (e.g., agências de viagem) que não têm um interesse num local específico, mas no processo de decisão do consumidor.
- *Covert Induced I* – Comunicação feita a partir de um promotor (embaixador) que se reconheça como popular no público-alvo e esteja associado ao destino turístico. Com estes esforços, as entidades esperam aumentar a credibilidade da publicidade, assim como a sua retenção.
- *Covert Induced II* – Promoção feita a partir de um promotor de viagem (escritor, *blogger*), que, aos olhos do recetor, não está associado a nenhuma entidade de turismo. Por sua vez, para o indivíduo, serve como uma fonte de informação não-enviesada (mais credível).
- *Autonomous* – Imagem passada a partir de filmes, notícias e artigos de jornal/revistas. Por norma, as entidades não têm qualquer controlo sobre a imagem transmitida a partir destes meios.
- *Unsolicited Organic* – trata-se de uma troca entre um indivíduo com experiência ou informação acerca de um destino e outro que pouco ou nada conhece acerca do local turístico.
- *Solicited Organic* – Uma das cinco etapas do comportamento do consumir é a busca por informação acerca dos produtos que se tem em seleção. Da mesma forma, aquando da seleção de um destino turístico, uma pessoa procura contrastar os pontos fortes e fracos. A informação orgânica solicitada é, portanto, toda a informação recebida por própria volição, informação que provém de outros indivíduos com mais experiência. O consumidor pretende confirmar as suas próprias crenças na seleção a partir de uma fonte fiável, que não seja enviesada pelo resultado. Assim, é possível assumir que a imagem orgânica é aquela que exerce uma maior influência na decisão de compra. Portanto, por ter uma influência tão grande, este evento social pode ser o momento que leva à decisão final, i. e., se haverá, ou não, compra (Litvin et al., 2008).
- *Organic* – A única fonte somente orgânica, onde o turista capta a imagem do destino na sua maior autenticidade, é a experiência. Gartner (1994) considera que este agente é a etapa final do continuum de formação da imagem de destino. Os turistas dotados com a experiência de viagem (imagem orgânica) tornam-se nos agentes de força orgânico solicitado e não-solicitado.

2.4. Imagem do destino online

Tradicionalmente, os estudos que se focaram na imagem do destino colocaram em hipótese aquilo que é comunicado pelas entidades que participam no mercado turístico e contrastaram estas comunicações com a percepção que o consumidor tem do destino (*e.g* Phelps, 1986). Não obstante, não será apenas a informação publicitária, que normalmente é enviesada, que participa no processo de interpretação de informação pela parte do consumidor. Existem agentes que induzem uma imagem do destino e agentes que proliferam uma imagem orgânica que está ilesa de interesse no resultado do evento comunicativo (Gartner, 1994).

Num mundo em que a tecnologia é acessível a uma base de utilizadores enorme, e as redes sociais fazem parte do dia-a-dia, surgem plataformas como o *TripAdvisor*, que servem especificamente para partilhar opiniões. Tendo em conta este fator, as entidades que participam no mercado turístico deixam de ser os maiores responsáveis pela construção de uma imagem do destino (Gedikoglu et al., 2020). Segundo Gedikoglu et al. (2020) os modelos para a conceptualização da imagem, em que se estuda a percepção dos turistas perante o destino, poderão estar desatualizados face ao contexto da globalização da tecnologia.

Os media sociais e a Internet, no geral, causaram uma mudança na forma como os consumidores adquirem produtos turísticos. Isto é, por um lado, os turistas fazem as suas decisões com base na informação presente no meio digital e na oportunidade de comunicação digital, o que torna os medias sociais numa competição direta aos meios de informação tradicionais (jornais, revistas, noticiários, companhias e agências de turismo). Por outro lado, a tecnologia diluiu a rede de intermediários, causando uma desintermediação: os turistas decidem e planeiam as suas viagens sem o auxílio dos intermediários tradicionais, como as agências de viagem (Amersdorffer et al., 2012).

Dwivedi (2009) conduziu um estudo que dá conta da imagem do destino online da Índia apoiando-se aos atributos da imagem percebida propostos em Beerli e Martín (2004, p. 659) para melhor compreender a imagem do destino percebida pelos turistas, e a partir destes atributos foi possível compreender e descrever a imagem do destino comunicada

no meio digital. Nos estudos de Dwivedi (2009) e Govers e Go (2008), verificou-se que os consumidores criam a sua própria imagem do destino.

Tendo em conta a natureza intangível da experiência turística (Litvin et al., 2008), um consumidor poderá estar constantemente à procura de nova informação, e ao adquirir toda esta informação, o consumidor irá criar uma imagem mental de como será o destino turístico a nível da experiência (Govers e Go, 2008).

Segundo Govers e Go (2008), o desenvolvimento da Internet e a respetiva disponibilização de novos multimédia com um maior grau de interatividade auxilia o consumidor a criar uma imagem percebida mais realística. Com efeito, é a partir da reunião de informação acerca do destino que o consumidor toma a sua decisão (Agapito et al., 2013; Gartner, 1994; Tasci e Gartner, 2007), e as plataformas disponíveis online apresentam-se como uma vasta fonte orgânica de informação (Dwivedi et al., 2009; Mak, 2017), o que pode afetar a intenção de visita pelos turistas (Mak, 2017).

O conteúdo criado por utilizadores disponível na Internet pode assumir vários formatos diferentes (Litvin et al., 2008). Neste sentido, e tendo em conta o modelo tridimensional de Gartner (1994), a imagem do destino online pode ser vista como a representação coletiva de sabedoria, crenças, sentimentos e interpretações globais de determinado destino. Diferindo, assim, da imagem do destino tradicional devido ao facto de, no digital, as perceções de cada turista são partilhadas, e, por isso, a imagem do destino online é resultado de uma cocriação entre os turistas (Mak, 2017). Ao passo que, segundo o ponto de vista tradicional da formação da imagem do destino, em que um sujeito reúne informação de fontes induzidas e orgânicas (Carvalho et al., 2020; Gartner, 1994; Phelps, 1986), no ciberespaço há uma maior facilidade em propagar e aceder a *word-of-mouth*, que é uma fonte de informação vista como mais credível do que comunicações vindas de esforços de marketing (Pan et al., 2007).

Goldsmith (2006) argumenta que a razão pela qual os consumidores partilham opiniões na Internet deve-se ao facto de, quando um indivíduo se interessa por determinada categoria de produto, este tira proveito de exibir a sua experiência e sabedoria acerca do produto. Os consumidores pesquisam por produtos na internet que, nem sempre, têm intenção de adquirir (Govers e Go, 2008). Por outro lado, os

2.5. Processamento de linguagem natural

O progresso da tecnologia e o sucessivo aparecimento da *Web 2.0* trouxe consigo inovações como as redes sociais e os *blogs* que permitem, facilmente, publicar conteúdos que ficam disponíveis a qualquer pessoa com acesso à Internet. Este tipo de conteúdo criado por utilizadores do ciberespaço atribui-se o nome de *user generated content* (UGC), ou, no âmbito do turismo, *tourist generated content* (TGC) (Mak, 2017; Marine-Roig e Huertas, 2020; Ren e Hong, 2017).

Em relação ao Marketing, este conteúdo é uma base de dados rica, visto que é constituída, em grande parte, a partir de testemunhos de consumo. Para o turismo, os *media* sociais tornaram-se numa fonte de informação importante para os turistas (Ren e Hong, 2017).

No que concerne à imagem do destino, o conteúdo elaborado pelos turistas na *web* é feito de forma arbitrária e sem interesse na influência que o conteúdo tem. Por isso, tornam-se em opiniões não enviesadas (Mak, 2017). O *UGC* é, portanto, um agente de formação de imagem orgânica, que pode ser solicitado ou não-solicitado (Gartner, 1994).

O *TGC* comporta-se, assim, como um veículo de *word-of-mouth* eletrónico que pode ser consultado em qualquer lugar do mundo e a qualquer hora. Por isso, tendo em conta a sua credibilidade, o conteúdo elaborado por turistas tem uma grande influência na formação da imagem do destino (Gartner, 1994; Jiang et al., 2021; Mak, 2017). No entanto, este conteúdo é imensamente vasto, e, por isso dificulta o tratamento e análise duma quantidade de dados exorbitante com métodos tradicionais (Ren e Hong, 2017), fazendo emergir a necessidade da sua automatização (Marine-Roig e Huertas, 2020).

Para facilitar a análise destes dados textuais é possível fazer uso de técnicas de processamento de linguagem natural, que permitem aos computadores compreender a linguagem humana (Marine-Roig e Huertas, 2020).

A imagem do destino forma-se a partir de pensamentos, sentimentos e motivações (Bigné et al., 2001; Echtner e Ritchie, 1991; Fakeye e Crompton, 1991; Gartner, 1994), ou seja, a partir de componentes de natureza abstrata. Neste sentido, a análise de

sentimento torna-se numa ferramenta extremamente útil na análise desta *Big Data* (TGC). Com esta técnica, é possível captar os sentimentos expressos nas opiniões dos turistas, ou seja, compreender de que maneira um destino é percebido na mente do consumidor (González-Rodríguez et al., 2016; Jiang et al., 2021).

Segundo Ortigosa et al. (2014) esta análise pode ser abordada de duas formas distintas:

I. É possível ter como base um léxico de palavras de opinião classificadas em diferentes categorias que dizem respeito à polaridade dessas palavras, e, com esta informação, será extraído o sentido semântico do texto para determinar se exprime um sentimento negativo ou positivo. Esta abordagem torna-se vantajosa quando o contexto é vago e é mais acessível devido à maior disponibilidade de bases lexicais.

II. Utilizar *machine-learning* com algoritmos de classificação (*supervised classification methods*), que têm como base um *training set* que distingue os diferentes polos ou categorias, de modo que o programa possa aprender a extrair essa informação. Ao contrário da abordagem anterior, esta servirá melhor para contextos específicos, como o turismo.

2.6. Análise de sentimentos em UGC (User Generated Content)

Desde Hunt (1975) é consensual que, para o sucesso do turismo de qualquer destino, estudos que exploram a imagem do destino são imperativos (Crompton, 1979; Gartner e Hunt, 1987). Não só no sucesso turístico em geral, como também no sucesso dos esforços de marketing e promoção (Calantone et al., 1989; Echtner e Ritchie, 1991; Fakeye e Crompton, 1991).

A imagem do destino possui impacto no comportamento do consumidor. Não só no *antes da viagem*, ou seja, no processo de tomada de decisão (Crompton, 1979; Graham., 1996; Hunt, 1975; Mayo, 1975) como no *durante* e no *depois* da viagem.

Um turista, antes de decidir o destino de viagem, baseia a sua escolha em várias fontes de informação, que ao mesmo tempo, atuam como agentes de influência (Gartner, 1994). Uma das fontes que possui mais peso nesta decisão é a imagem do destino percebida. Algo que é fundamental durante o processo de escolha/compra, e que também terá uma

grande influência na satisfação do indivíduo durante e após finalizar a viagem, devido à expectativa que é criada com base na imagem (Bigné et al., 2001; Chon, 1992). Com efeito, quanto mais a imagem se afastar da experiência real, maior será a insatisfação do turista no fim da viagem (Phelps, 1986), e conseqüentemente, na sua intenção de não voltar a fazer a viagem (Bigné et al., 2001).

Por isso, a imagem pode ser considerada uma representação do local turístico em que, por sua vez, as diferentes componentes e dimensões da imagem influenciam o processo de tomada de decisão (Tasci e Gartner, 2007).

Do ponto de vista do Marketing, a compreensão da imagem que um destino tem, em termos dos pontos fortes e fracos dos destinos turísticos, e se a imagem passada é realista, é necessária para desenvolver estratégias eficientes (Echtner e Ritchie, 1991). Criar e gerir uma boa imagem é uma componente fulcral para se posicionar corretamente na mente do consumidor e atrair o público-alvo desejado (Calantone et al., 1989). A publicidade é imperiosa no turismo devido à sua intangibilidade (Fakeye e Crompton, 1991).

Os conteúdos elaborados pelos turistas (*TGC*) comportam-se como uma atividade de *word-of-mouth* no espaço digital (*eWoM*) (Mak, 2017). Litvin et al. (2008, p. 458) argumenta que “o princípio fundamental do comportamento do consumidor é a habilidade de exercer influência sobre outros”, neste sentido, é possível afirmar a importância que a recomendação de um produto a outros consumidores tem na sua decisão final (Jalilvand et al., 2012).

Quando um turista expressa a sua opinião é criada uma oportunidade para outros viajantes que ainda não viveram a mesma experiência, a conhecerem os atributos mais relevantes de determinado destino. Esta conversa digital, por influenciar a imagem do destino, tem um impacto significativo no processo de decisão de viagem (Gartner, 1994; Jalilvand et al., 2012). Um bom aproveitamento dos atributos dum destino (e.g., hotéis, restaurantes ou entretenimento) resulta na propagação de *eWoM* positivo. Por isso, conseguir reconhecer, especificamente, quais os atributos de um destino que suscitam um maior aproveitamento é importante para o desenvolvimento turístico (Jalilvand et al., 2012).

Segundo Nowacki e Niezgodá (2020), a partir de técnicas de *Natural Language Processing*, como *text mining*, é possível identificar as palavras mais frequentes entre a amostra de opiniões. Assim, possibilita-se o reconhecimento do que é considerado como atributos de determinado destino pelos turistas, o que diz respeito à componente cognitiva da imagem (Gartner, 1994). A partir da análise de sentimentos consegue-se compreender a avaliação, em termos de polaridade (negativa ou positiva), que os turistas fazem aos destinos, o que resulta em informação sobre a componente afetiva da imagem (Gartner, 1994).

A análise de sentimento fornece informação acerca dos sentimentos expressos pelos atributos do destino, o que contribui na criação de melhores estratégias de Marketing, e, conseqüentemente, aumentam a lealdade dos turistas a partir de uma melhor oferta no turismo (Nowacki e Niezgodá, 2020).

2.6.1. Implementação da análise de sentimentos

A análise de sentimentos é aplicada a uma base de dados qualitativos, que possui as opiniões dos indivíduos. O mais comum, é esta amostra ser retirada da Internet (Ortigosa et al., 2014), por isso, para conduzir uma análise de sentimentos, é necessário, em primeiro lugar, capturar as opiniões a partir de plataformas sociais, o que pode ser feito manualmente, ou automatizado com a utilização de *webcrawlers* (El-Masri et al., 2017; Jiang et al., 2021; Liu et al., 2020; Ren e Hong, 2017).

Após a extração dos textos é necessário efetuar uma limpeza aos dados, fase denominada de pré processamento. Existem diversas de técnicas que são utilizadas com o intuito de auxiliar o seu processamento algorítmico (NLP), aumentando a eficácia da análise e a precisão dos resultados (Altrabsheh et al., 2014), e este processo trata-se da remoção de elementos desnecessários na amostra (El-Masri et al., 2017). El-Masri et al. (2017) destaca as seguintes técnicas: *tokenization*, que separa o texto em unidades mais reduzidas; *stopwords*, que elimina palavras irrelevantes, que são palavras que não expressam qualquer tipo de sentimento, e que, por isso, não contribuem para a polaridade; remoção da pontuação e de números, visto que não expressam polaridade de sentimentos e, por isso, não contribuem para a análise; alteração de todo o texto para letras minúsculas

(ou maiúsculas) com o objetivo de unificá-lo e ajustá-lo ao *training set*; e, por último, correção ortográfica.

Ademais, é necessário definir o processo de identificação ou discriminação, de categorias linguísticas. Isto é, ensinar ao computador, ou impor-lhe regras, da forma como texto deve ser interpretado. Existem diversos parâmetros na análise de sentimento que podem ser utilizados para este fim, no entanto, segundo El-Masri et al. (2017), os mais comuns são:

- (I) os *n*-gramas, onde se define quantas palavras são interpretadas de cada vez (Go et al., 2009);
- (II) e a categorização gramatical (*part-of-speech tagging*), que consiste na classificação das palavras em classes de palavras.

Após todos os parâmetros definidos, a análise pode ser executada. Desta forma procede-se à escolha do método com que se realiza a análise de sentimento (base lexical ou inteligência artificial).

Em suma, esta técnica de análise tem sido muito aplicada no turismo (González-Rodríguez et al., 2016; Liu et al., 2020; Park et al., 2020; Ren e Hong, 2017), assim como nas subcategorias que integram o ecossistema de serviços do sector, para aumentar o conhecimento que se tem sobre a imagem de determinados destinos. Muitos estudos procuram captar os atributos dos destinos e a sua respetiva avaliação (Huang et al., 2020; Lalicic et al., 2021; Nowacki e Niezgodna, 2020), outros focam-se em estudar o comportamento do consumidor (González-Rodríguez et al., 2016; Park et al., 2020b; Yan et al., 2015), verificar alterações na imagem ao longo de um período de tempo (Liu et al., 2020; Marine-Roig e Huertas, 2020). Não obstante, é possível simplesmente medir, ou fazer um reconhecimento geral acerca da imagem de um destino (Jiang et al., 2021), sendo também possível aplicar a análise de sentimento em áreas específicas, como os serviços que fazem parte da oferta turística, por instância, a hotelaria e restauração (Ye et al., 2011; Yu e Zhang, 2020).

2.7. Presente estudo

Durante a última década surgiram inúmeros estudos em que os conceitos teóricos da imagem do destino (Agapito et al., 2013; Beerli e Martín, 2004; Bigné et al., 2001; Gartner, 1994; Hunt, 1975) são aplicados a dados textuais não estruturados extraídos da Internet (UGC), a partir dos quais se efetua uma análise de sentimento para, de um modo global, compreender a imagem do destino percebida pelos turistas de determinado destino ou serviços acessíveis no destino (Huang et al., 2020; Jiang et al., 2021; Marine-Roig e Huertas, 2020; Park et al., 2020; Ren e Hong, 2017; Yu e Zhang, 2020).

Tipicamente, o objetivo destes estudos é transformar o que é comunicado pelos utilizadores em dados quantitativos estruturados, o que é feito por meio de uma análise de conteúdo, que atenta agrupar os termos que partilham semelhanças em categorias exclusivas. Este objetivo pode ser cumprido a partir de uma análise de frequência que assenta no princípio de que apenas os termos mais frequentes são aqueles que revelam mais interesse (Marine-Roig e Huertas, 2020). Em Jiang et al. (2021) é elaborado um estudo para compreender a imagem do destino online que os turistas vindos da China têm de Hong Kong, e avalia-la ao longo de um período entre 2012 e 2018. É, precisamente, a partir da frequência de palavras (juntamente com a valência emocional) que foi possível fazer sentido do conteúdo textual e enquadrá-lo nas dimensões da imagem do destino (cognitivo-afetivo). Neste estudo (Jiang et al., 2021), a preparação dos dados para a análise seguiu o seguinte procedimento:

I. Extração dos dados e limpeza da base de dados: com efeito, é necessário remover informação desnecessária dos dados para torná-los mais consistentes (Altrabsheh et al., 2014; El-Masri et al., 2017).

II. Extração de categorias: os dados textuais têm de passar por um processo de segmentação lexical, onde se discrimina cada classe de palavra (*e.g* nome, verbo, pronome e adjetivo), este passo é fundamental para que seja possível executar a análise de sentimento.

III. Análise de sentimento: a valência emocional é obtida por meio de um modelo de *deep learning*, que faz uso de dados que já estão classificados quanto ao sentimento, para treinar e prever os resultados dos dados textuais que lhe são submetidos.

Tendo os dados preparados e os resultados da análise de sentimento obtidos, é possível enquadrar o conteúdo (UGC) no teorema da imagem do destino. Assim, a partir da frequência com que as características (atrações, atividades, meio-ambiente e pessoas) de Hong Kong são evocadas nos comentários dos turistas, e a respetiva avaliação positiva-negativa, é possível auferir quanto à dimensão cognitiva da imagem do destino. A classificação das palavras executada durante a extração de categorias permite descobrir quais os adjetivos mais usados. Estas palavras exprimem sentimentos e estados de mente (Marine-Roig e Huertas, 2020) que podem ser positivos ou negativos, o que ajuda na compreensão das emoções adjacentes aos comentários dos turistas, revelando a imagem afetiva que os turistas chineses têm de Hong Kong (Jiang et al., 2021).

Há um consenso no procedimento adotado em Jiang et al. (2021) em relação aos vários estudos (Huang et al., 2020; Lalicic et al., 2021; Marine-Roig e Huertas, 2020; Yu e Zhang, 2020) que utilizam a análise de sentimento para investigar a imagem de diferentes destinos. Isto é, é comum que as componentes cognitivas da imagem do destino sejam agrupadas em categorias gerais de modo a compreender quais são os pontos em que o destino produz maior e menor satisfação, e que este processo é feito tendo em conta a frequência com que as palavras são utilizadas entre todas as opiniões de turistas. Igualmente, os adjetivos exprimem emoções negativas ou positivas e, por sua vez, referem-se a características funcionais do destino. A dimensão conativa da imagem pode ser reconhecida no conteúdo a partir de termos que exprimam recomendação positiva e negativa (Marine-Roig e Huertas, 2020).

De modo a revelar a imagem percebida do destino, em Beerli e Martín (2004) é compilado uma série de fatores que influenciam a percepção que os turistas desenvolvem da imagem dos destinos, esta lista compreende as nove dimensões (recursos naturais; lazer e recreação turística; meio-ambiente natural; infraestruturas gerais; infraestruturas turísticas; cultura, história e arte; meio-ambiente social; fatores político-económicos; atmosfera geral do destino) com os respetivos atributos. Como se pode verificar em Liu et al. (2020), Jiang et al. (2021) e Dwivedi (2009), aliar estas dimensões e atributos à análise de sentimento possibilita tirar conclusões pertinentes acerca da imagem do destino. Por instância, é possível perceber, dentro de cada dimensão, quais os aspetos mais marcantes no destino. No caso de Liu et al. (2020), leva à conclusão de que os turistas internacionais de Macau são mais motivados para visitar este destino devido à sua

herança histórica e cultural, e devido à sua oferta de lazer e recreação. Em relação ao estudo de Jiang et al. (2021), foi possível compreender de que maneira as atrações, as atividades, o meio-ambiente e as pessoas são percebidas e qual a opinião dos turistas sobre estas categorias dentro de um determinado período temporal, o que demonstra objetivamente a performance dos diferentes componentes do destino ao longo do tempo e por que razão melhoraram ou pioraram. Igualmente, em Dwivedi (2009), ao analisar o conteúdo textual e compreendê-lo nas nove dimensões, possibilitou conhecer os tópicos mais discutidos pelos turistas da Índia, e assim perceber que locais turísticos são mais consagrados e as respetivas preocupações dos consumidores.

No presente estudo, o investigador irá aplicar os conceitos teóricos da imagem do destino, nomeadamente, o modelo tridimensional da imagem do destino de Gartner (1994) que compreende a imagem em três elementos hierárquicos (cognitivo – afetivo – conativo) devido ao facto de ser largamente apoiado e está, efetivamente, demonstrado a sua utilidade na compreensão da imagem do destino (Agapito et al., 2013; Beerli e Martín, 2004; Bigné et al., 2001; Carvalho et al., 2020; Echtner e Ritchie, 1991). Tendo em conta que a amostra do estudo é proveniente do conteúdo publicado por turistas, far-se-á uso da *framework*, apresentado na Figura 2.1, de Echtner e Ritchie (1991) que possibilita medir a imagem do destino em três eixos diferentes (Atributos – holístico; características funcionais – características psicológicas; comum – único), com esta ferramenta pretende-se visualizar quais características do Algarve são consideradas comuns ou únicas pelos turistas. Para ter uma noção de qual é a imagem cognitiva percebida pelos consumidores de produtos turísticos do Algarve, os dados serão organizados segundo as dimensões propostas por Jiang et al. (2021), que são adaptadas a partir das categorias propostas por Beerli e Martín (2004). Com efeito, o primeiro objetivo que este estudo pretende cumprir é a descrição detalhada da imagem do Algarve enquanto região turística, que será auxiliada pelos modelos supracitados.

É evidente na literatura (Baloglu e McCleary, 1999; Cardoso et al., 2019; Stylos et al., 2016; Tasci e Gartner, 2007; Zhang et al., 2014) que a dimensão cognitiva exerce influência na dimensão afetiva, que por sua vez age como intermediário entre a cognição e a atitude final (dimensão conativa) (Agapito et al., 2013; Papadimitriou et al., 2018). Tendo em conta este fator, este estudo pretende, também, averiguar a relação existente

entre a imagem cognitiva e afetiva do Algarve, e o efeito consequente na sua imagem global e comportamento do consumidor posteriormente à experiência.

Vários autores demonstraram a eficiência de métodos de análise assistidos por aprendizagem automática (Dwivedi, 2009; Jiang et al., 2021; Lalicic et al., 2021; Liu et al., 2020; Marine-Roig e Huertas, 2020; Ren e Hong, 2017). Estes métodos são especialmente úteis no tratamento de bases de Mega dados (Marine-Roig e Huertas, 2020). Partindo da premissa de que a análise de sentimento dispõe a oportunidade de observar a valência emocional presente nas opiniões dos turistas (Jiang et al., 2021), e sucessivamente tirar conclusões sobre a satisfação dos turistas (Yu e Zhang, 2020), propõe-se avaliar a satisfação dos turistas a partir deste método, assim como discriminar os produtos/serviços mais e menos consagrados pelos visitantes no Algarve.

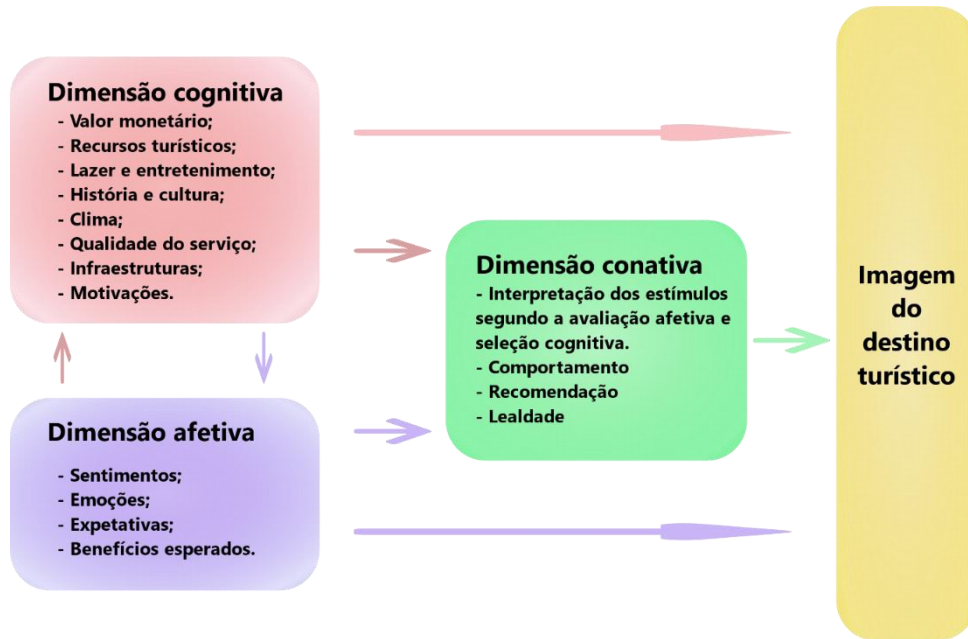
O presente estudo pretende alcançar os objetivos estipulados a partir de uma análise quantitativa, que parte de uma base de dados constituída por opiniões deixadas por antigos consumidores internacionais de produtos e serviços turísticos no Algarve. Até à data, não há estudos que procuram compreender a imagem do Algarve enquanto destino turístico a partir de uma análise de sentimento, tornando o presente estudo no pioneiro deste método para analisar a opinião dos turistas internacionais e aferir o desempenho deste destino. Para tal, a investigação terá como metodologia a recolha, limpeza de base de dados, extração de classes gramaticais, processamento da análise de sentimento e interpretação de dados. Por fim, os dados serão tratados com uma abordagem estatística de modo a obter informação relevante e, também, visualizá-los mais facilmente.

2.8. Modelo conceptual

Gartner (1994) distingue a imagem do destino em três dimensões (cognitivo – afetivo – conativo). Segundo a literatura (Agapito et al., 2013; Bigné et al., 2001; Gartner, 1994), as dimensões interagem entre si, sendo que a relação entre a dimensão cognitiva e afetiva resulta na imagem global. Partindo desta premissa, este estudo adaptou o modelo de investigação presente em Jiang et al. (2021), adaptando-o ao destino em questão no estudo (Fig. 4). Este modelo compreende a imagem em quatro categorias diferentes (atrações, atividades, meio-ambiente e pessoas), e em cada categoria encontram-se subcategorias que especificam as características presente nas opiniões. Adicionalmente, este modelo

auxilia na compreensão da imagem do destino a partir da análise de sentimento, - as revisões, uma vez categorizadas, são sujeitas à valência emocional e posteriormente organizadas e filtradas em dimensões (cognitiva – afetiva). O modelo conceptual adotado para esta dissertação é ilustrado pela figura 3:

Figura 2.3. Modelo conceptual da formação da imagem do destino turístico.



Fonte: adaptado de Peña et al. (2012: 263)

De modo a obter uma descrição holística da imagem do Algarve, a dimensão conativa é aferida a partir do método utilizado por Marine-Roig e Huertas (2020), em que se mede o comportamento posterior à experiência a partir de termos relacionados com a recomendação positiva ou negativa. Segundo o estudo (Marine-Roig e Huertas, 2020), a dimensão afetiva pode, também, ser medida a partir da frequência de termos relacionados com sentimentos e estados de espírito positivos e negativos.

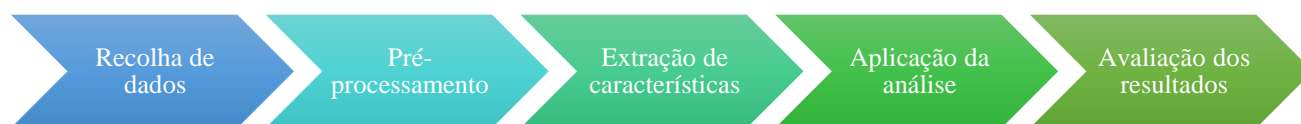
3. METODOLOGIA DA INVESTIGAÇÃO

3.1. Aplicação do método

Tendo em conta a natureza do presente estudo, isto é, a conversão de conteúdo textual em dados e a respetiva análise assistida por inteligência artificial, surge a necessidade de aplicar métodos de NLP (*Natural language processing*) aos dados extraídos do *TripAdvisor*, com o objetivo de efetuar uma análise de sentimento.

Com o objetivo de obter o devido sucesso, esta pesquisa adota a seguinte metodologia, sugerida em El-Masri et al. (2017):

Figura 3.1. Estrutura da metodologia para o processamento de linguagem natural



Fonte: Adaptado de El-Masri et al. (2017: 54)

3.1.1 Recolha de Dados

A base de dados criada em prol do presente estudo é proveniente da plataforma online *TripAdvisor*, devido ao facto de a mesma ser a maior plataforma de viagens do mundo, contanto com mais de 934 milhões opiniões de turistas e mais de 8 milhões de produtos turísticos envolvidos. Esta plataforma permite filtrar os diferentes produtos em diferentes categorias, nomeadamente, alojamento, atividades e restauração. De modo a obter os dados, foram selecionados diferentes hotéis, restaurantes, atividades e atrações em diferentes regiões do Algarve com o objetivo de construir uma base com uma quantidade equilibrada de instâncias entre os diferentes produtos e serviços turísticos, e com o apoio do software *WebHarvy* foi possível extrair cada opinião. Assim, utilizando os filtros da plataforma, foi possível selecionar apenas os locais mais populares entre os turistas. Com efeito, para o presente estudo extraiu-se 2923 opiniões relacionadas com atividades e atrações turísticas, e após rever a amostra removeu-se opiniões incompletas e/ou

incongruentes, resultando num total de 2847 instâncias. Os dados obtidos nesta fase são testemunhos orgânicos (Gartner, 1994) de turistas internacionais.

3.1.2. Pré-processamento

De modo a construir uma base de dados que produza resultados precisos na análise de sentimentos (Altrabsheh et al., 2014) é necessário remover elementos desnecessários ao processamento de informação, as técnicas de pré-processamento, selecionadas segundo o estudo de El-Masri et al. (2017):

- Tokenizar: Purgar toda a pontuação no conteúdo textual.
- Transformar: Todo o conteúdo é convertido em letra minúscula e todas as hiperligações são removidas.
- Filtrar: Uma lista de palavras não relevantes (*e.g* artigos, pronomes e conjunções) foi utilizada. Estas palavras são omitidas para evitar que, por serem as mais comuns na língua inglesa, interferissem.

Nesta fase do tratamento dos dados seleciona-se a sequência de *n-gramas* pretendida, ao definir este valor é também estipulado de que forma o programa interpreta o texto, aumentando, também, o desempenho da análise. Assim, no presente estudo foi decidido considerar a combinação de unigramas ($n=1$) e bigramas ($n=2$) (El-Masri et al., 2017). Esta técnica é útil na observação de combinações de palavras que expressam significados únicos (*e.g. well worth e good value*) e que surgem frequentemente no conteúdo textual.

Com o fim de evitar ambiguidades no texto, é possível fornecer ao programa informação acerca das classes gramaticais utilizando um modelo *part-of-speech tagging* (*averaged perceptron tagger*), que classifica todas as palavras na sua classe gramatical. O resultado desta classificação permite filtrar diferentes classes de palavras com o fim de extrair adjetivos com valor emocional e atributos cognitivos a partir da classe dos nomes.

Por fim, é necessário converter o conteúdo textual em informação que a inteligência artificial consiga compreender, o que pode ser conseguido a partir de um algoritmo que transforma o texto num vetor de palavras, neste estudo aplica-se o método TF-IDF (*term frequency – inverse document frequency*), que calcula a relevância, ou mede o peso, que

determinado termo tem em toda a base de dados excluindo as palavras comuns e considerando apenas a frequência dos termos mais raros (Shah et al., 2020). A fórmula do algoritmo é a seguinte (Shah et al., 2020):

$$TFIDF(w) = TF(w) \times IDF(w)$$

(1)

TF – *Term Frequency*, frequência com que o termo (w) surge em determinado documento:

$$tf(w) = \log(1 + freq(w))$$

(2)

IDF – *Inverse Document Frequency*, calculo que deduz, para determinado termo (w), se é raro ou comum:

$$idf(w) = \log\left(\frac{N}{count(d \in D: w \in d)}\right)$$

(3)

Nesta formula, N é o número total de documentos, e a expressão $count(d \in D: w \in d)$ significa o número total de documentos em que o termo (w) surge (Paik, 2013).

Esta técnica é, também, útil na extração de palavras-chave presentes em todas as opiniões, e partindo da premissa de que a relevância de cada termo exprime os aspetos mais importantes no conteúdo criado por turistas é possível obter informação valiosa acerca das dimensões da imagem do destino.

Segundo Marine-Roig e Huertas (2020), a análise de conteúdo é uma técnica que pode ser utilizada com o objetivo de elaborar uma descrição quantitativa e sistemática do conteúdo. Para tal, é necessário criar categorias que permitam agrupar os dados em grupos que partilham semelhanças. O presente estudo aplicou um método *à priori* para categorizar os dados, isto é, os dados passaram por um processo de categorização manual que agrupa as opiniões em quatro categorias diferentes, em que cada uma se divide em diferentes subcategorias:

A. Atrações

A1. Naturais

A2. Humanas

A3. Históricas e culturais

A4. Costumes populares

B. Atividades

B1. Alojamento

B2. Comida e restauração

B3. Comércio local

B4. Lazer e entretenimento

C. Ambiente

C1. Saúde pública

C2. Preço

C3. Clima

C4. Infraestrutura e atmosfera

D. Pessoas

D1. Prestadores de serviço

D2. Residentes locais

Esta categorização foi inspirada e adaptada a partir do método presente em Jiang et al. (2021) devido à sua aplicabilidade no Algarve enquanto destino. Tendo em conta que o estudo que apresentou esta categorização tinha como objetivo investigar um destino que não o Algarve, foram suprimidas categorias do original para melhor refletir a realidade.

3.1.3. Extração de características

Após a fase de limpeza e pré-processamento do conteúdo textual, torna-se possível a extração das variáveis que permitem estruturar e explorar a imagem do destino. As características extraídas da amostra foram as seguintes:

- *N-gramas*: durante o pré-processamento, é aplicada uma transformação que converte cada palavra ($n=1$) ou cada conjunto de palavras ($n=2$) em itens que podem, posteriormente, ser sujeitos a diferentes métricas.
- *Frequência de palavras*: o método TF-IDF permite obter a frequência com que determinada palavra surge dentro do número total de palavras em toda a amostra. Esta característica permite visualizar, em termos gerais, as palavras mais significativas na sintetização da opinião partilhada dos turistas.
- *Categorias gramaticais*: A classificação gramatical (*POS Tagging*) atribui, automaticamente, uma etiquetagem que discrimina todas as palavras em classes gramaticais.
- *Valência*: a análise de sentimentos permite traçar a média do sentimento expresso na opinião partilhada dos turistas, o que permite segmentar o conteúdo textual em duas grandes categorias, as críticas positivas e as críticas negativas.
- *Palavras-chave positiva e negativa*: a junção entre a classificação gramatical e o método TF-IDF possibilita a extração de palavras que implicam uma avaliação em cada categoria.

3.1.4. Aplicação do algoritmo da análise de sentimentos

De modo a analisar, fazer sentido deste conteúdo textual e tirar conclusões acerca da avaliação feita pelos consumidores do destino turísticos, estudos anteriores procederam à utilização da análise de sentimentos para obter conclusões acerca da imagem do destino (Huang et al., 2020; Lalicic et al., 2021; Marine-Roig e Huertas, 2020; Nowacki e

Niezgoda, 2020; Park et al., 2020; Yu e Zhang, 2020). Assim, esta técnica serve para identificar e classificar sentimentos expressos no discurso, de modo a polarizá-los com o objetivo de interpretá-los mais facilmente (González-Rodríguez et al., 2016).

Para conduzir este tipo de análise, durante esta investigação fez-se uso do software *Orange Data Mining*, versão 3.29, que se trata de uma *graphic user interface* (GUI) que executa códigos em Python. Esta ferramenta permite executar, entre outros modelos de valência

A análise de sentimento, no *Orange Data Mining*, pode ser operacionalizada a partir de diferentes modelos de classificação sentimental. No entanto, antes da análise, considerou-se os seguintes algoritmos:

1. VADER (*Valence Aware Dictionary and sEntiment Reasoner*): o algoritmo VADER é um modelo baseado em regras parcimoniosas, focado na classificação/análise de sentimentos implícitos em textos provenientes das redes sociais. Este modelo surge devido à dificuldade de aplicar a análise de sentimentos em contexto de *microblogging* devido ao enorme volume de dados. Em Hutto e Gilbert (2014) é apresentado um algoritmo que pretende solucionar este problema. Este modelo opera segundo um léxico manufaturado que inclui uma métrica que atribui a valência expressa por determinado texto. Curiosamente, este algoritmo demonstra uma precisão acima da precisão humana, - em Hutto e Gilbert (2014), ao comparar o grau de precisão deste modelo com a precisão humana, é demonstrado que a métrica de avaliação *F1* (precisão) tem um valor de 0.96, superando a precisão da classificação humana ($F1 = 0.84$).

2. Hu-Liu04: em Hu e Liu (2004) é demonstrada uma outra forma de resolver o problema da análise de sentimento em conteúdo textual, resolvendo-o a partir de um método que, em vez de processar os dados textuais segundo um léxico de palavras, a integração da análise opera segundo uma extração extensa de categorias. Segundo Hu e Liu (2004), esta análise de sentimento recorre a (1) extração de categorias gramaticais (*POS Tagging*); (2) extração de frases que ocorrem com frequência; (3) exclusão de características não-genuínas, analisando a distância entre as palavras de cada frase extraída; (4) extração de palavras de opinião (*opinion words*), identificando e extraíndo adjetivos adjacentes a características; (5) identificação de características

excepcionais, que é obtida a partir das palavras de opinião. No caso de determinada frase não conter nenhuma característica frequente, no entanto, contém uma ou mais palavras de opinião, o nome mais próximo a essa palavra é extraído como uma característica excepcional. A precisão deste método é medida durante o estudo (Hu e Liu, 2004) em comparação à precisão humana, desempenhando com uma precisão de 0.72.

Comparando a precisão do desempenho dos diferentes algoritmos, o método VADER (Hutto e Gilbert, 2014) foi escolhido devido ao facto de que demonstra uma maior precisão e, conseqüentemente, torna os resultados mais próximos da realidade.

3.1.5. Avaliação dos resultados

Tendo aplicado as técnicas de pré-processamento do texto, extraído as características relevantes para a análise, e, tendo obtido o resultado da análise de sentimentos, é possível avaliar os resultados e enquadrá-los no problema que o presente estudo pretende resolver. O próximo capítulo desta dissertação ilustrará a informação obtida após a aplicação a metodologia sugerida em El-Masri et al. (2017).

4. ANÁLISE DE RESULTADOS

Aproveitando as características extraídas a partir do pré-processamento e análise de sentimentos é possível proceder à análise dos dados.

As categorias são úteis na segmentação do texto, que em conjunto com a valência e frequência de palavras resulta num resumo de cada um dos fatores que constitui a imagem do destino. As etiquetas *POS* filtram classes de palavras, o que resume atributos e sentimentos subjacentes às críticas, proporcionando listas de termos que expressam um significado global em relação à amostra.

Os dados recolhidos designam-se de conteúdo textual não estruturado. Estes dados foram extraídos da plataforma social *TripAdvisor*, que se trata de uma plataforma social online dedicada ao turismo e para os turistas, que permite aos seus utilizadores partilhar opiniões, e avaliar os produtos ou serviços em diferentes locais, por eles visitados. A extração foi possível a partir do software *WebHarvy*, que possibilita a automatização de tarefas, tais como, recolher texto na Internet. Neste caso, foi possível recolher críticas, que em conjunto exprimem um significado partilhado entre todos os utilizados. Os dados recolhidos correspondem ao período entre 2005 e até à presente data.

4.1. Imagem cognitiva

Primariamente, e tendo em conta a natureza hierárquica da imagem do destino, pretende-se ilustrar a dimensão cognitiva do Algarve enquanto destino turístico. Assentando na premissa de que a dimensão cognitiva engloba os atributos físicos, palpáveis e mensuráveis presentes no destino (Echtner e Ritchie, 1991; Hernández-Mogollón et al., 2014), esta tabela ilustra a imagem cognitiva em grupos de categorias de diferentes. Para cada categoria segue a média da pontuação de valência obtida na análise de sentimento VADER; a percentagem, dentro do número total de críticas, com que a categoria é evocada no conteúdo; a quantidade total de críticas positivas e negativas, seguidas do rácio de críticas positivas dentro de cada categoria; por fim, a partir da frequência de palavras retirou-se, manualmente, as palavras-chaves das críticas positivas

e negativas. A partir da Tabela 4.1 pretende-se explorar a imagem cognitiva do Algarve, e realçam-se as seguintes impressões:

Tabela 4.1. Média da valência das categorias e análise de palavras-chave por categoria.
Adaptado de Jiang et al. (2021: 11)

Grupo	Categoria	Valência	Nº de críticas	% do total	Críticas positivas	Críticas negativas	Rácio das críticas positivas	Palavras-chave positivas	Palavras-chave negativas
A - Atrações	Naturais	0.78	1198	14.2%	1147	51	95.7%	beautiful, clean, big	poor
	Humanas	0.84	370	4.4%	361	9	97.6%	interesting, well worth	degraded
	Históricas e Culturais	0.86	301	3.6%	295	6	98.0%	fantastic, old	
	Costumes populares	0.89	39	0.5%	39	0	100.0%	authentic, different	
B - Atividades	Alojamento	0.75	543	6.4%	491	52	90.4%	great, clean, comfortable	old, basic, small, dirty
	Comida e restauração	0.84	1122	13.3%	1076	46	95.9%	friendly staff	cheap, small, awful
	Comércio local	0.81	55	0.7%	53	2	96.3%	many, local, worth	
	Lazer e entretenimento	0.83	975	11.6%	933	42	95.7%	worth, fun, wonderful	worst, avoid
C - Ambiente	Saúde Pública	0.83	751	8.9%	718	33	95.6%	clean, safe	broken, dirty
	Preço	0.82	506	6.0%	478	28	94.5%	good value	expensive, average, bad
	Clima	0.81	46	0.5%	43	3	93.5%	sunny, warm, hot	
	Infraestrutura e atmosfera	0.79	815	9.7%	777	38	95.3%	quiet, plenty, easy	crowded, noisy
D - Pessoas	Prestadores de serviços	0.76	1645	19.5%	1561	84	95.0%	helpful, friendly	busy, cold, rude
	Residentes	0.79	67	0.8%	63	4	94.0%	lovely, friendly	

As atrações naturais, categoria que inclui a praia, o mar, as reservas naturais, e a paisagem da serra, é a segunda categoria mais evocada nas opiniões dos turistas (14.2%). Tendo em conta que o turismo no Algarve se foca nos seus recursos naturais, a popularidade deste atributo é expectável. Os termos mais relevantes que definem esta

categoria, segundo as opiniões, é *beautiful* (lindo), *clean* (limpo) e *big* (grande) em críticas positivas, e *poor* em críticas negativas. As palavras-chave positivas refletem a dimensão da costa algarvia, e limpeza das praias e a beleza observada no meio-ambiente. Por outro lado, a palavra-chave que define os comentários negativos reflete a experiência que certas atrações naturais criam, quando estas estão em más condições. A valência desta categoria tem uma média de 0.78, o que demonstra que há uma grande satisfação por estes atributos.

As atrações humanas apresentam uma procura mais reduzida, observando-se apenas 4.4% de frequência com que são referidas em todas as críticas. No entanto, demonstra que é uma característica apreciada no Algarve a partir da valência (0.84) e o alto rácio de críticas positivas. Observando as palavras-chave positivas e negativas, os turistas apreciam o interesse (*interesting*) que é suscitado por estas atrações e consideram que é tempo e/ou dinheiro bem empregue (*well worth*). Por outro lado, algumas destas atrações encontram-se degradadas (*degraded*).

As atrações históricas e culturais parecem, assim como as atrações humanas, pouco procuradas tendo em conta o valor percentual (4.4%) com que esta categoria é evocada nas críticas. No entanto, a valência é das mais altas (0.86), o que pode dever-se ao facto de haver poucas referências nas críticas dos turistas. De acordo com as palavras-chave, a história e cultura de Portugal é positivamente descrita como antiga (*old*) e fantástica (*fantastic*) no sentido de se apresentar dum modo imersivo e distanciar o indivíduo da realidade.

A categoria “costumes populares” é a menos frequente nas críticas (0.5% do total). Estes costumes podem ser desconhecidos por muitos turistas devido ao facto de que há pouca exposição deste elemento. Esta categoria inclui festividades medievais no castelo de Castro Marim e tradições típicas na serra algarvia. Verifica-se que, esta categoria, apesar de pouco frequente, demonstra uma valência altamente positiva (0.89), sem qualquer crítica negativa. As palavras-chave revelam que as tradições típicas no algarve são autênticas/genuínas (*authentic*) e pouco comuns (*different*).

Relativamente aos serviços de alojamento, os turistas procuram espaços limpos (*clean*) e confortáveis (*comfortable*), e evitam voltar a lugares com um aspeto antigo (*old*), básico

(*basic*) e sujo (*dirty*). Preferencialmente, as instalações devem ser espaçosas, devido à palavra-chave negativa *small* (pequeno). O alojamento é uma categoria que não é tão frequente, sendo evocada em apenas 6.4% das críticas, e é a que tem um rácio de críticas positivas mais baixo (90.4%) com uma média de valência de 0.75.

Os serviços de comida e restauração presentes no Algarve é uma categoria evocada em 13.3% do total de críticas, o que demonstra a importância que a disponibilidade destes estabelecimentos tem para o turista. É uma categoria bastante prezada no destino, sendo que tem uma valência de 0.84, e um rácio de 95.9% de críticas positivas. A palavra-chave positiva “*staff*” indica o papel fulcral que os provedores de serviço têm neste serviço. Por outro lado, as palavras-chave negativas referem-se ao produto, ou seja, a comida. As palavras *cheap* (barato) e *awful* (horrível) indica que a qualidade da comida e do serviço exerce um peso na valência, assim como a dimensão do espaço e/ou quantidade de comida (*small* – pequeno).

O comércio local é a categoria menos procurada pelos turistas, demonstrando apenas 0.7% de evocação nas críticas analisadas. No entanto, os consumidores consideram que há variedade de estabelecimentos de comércio local (*many, local*) e que os produtos que se pode adquirir estabelecem uma relação de valor (*worth*).

Existe uma forte procura de fontes de lazer e entretenimento no Algarve, fundamentado pela percentagem de evocação desta categoria (11.6%). Este atributo é, também, uma forma de experiência agradável, resultando numa valência de 0.83. Esta categoria inclui os parques aquáticos, as rotas turísticas por mar e por terra e as atividades marítimo-turísticas.

A higiene e saúde pública é um fator importante na experiência turística. No caso do Algarve, esta categoria está presente em 26.4% das críticas, com uma valência de 0.83. Note-se que, a higiene e segurança, é um interesse revelado noutros atributos da imagem devido à presença de palavras-chave relacionadas com a saúde pública. Os turistas apreciam a segurança e limpeza (*clean, safe, broken, dirty*)

O valor monetário (categoria “preço”) pago pelos bens e serviços no Algarve é um atributo evocado em 8.9% das críticas. A valência desta categoria (0.82) demonstra que,

no geral, os preços praticados no Algarve são razoáveis. As palavras-chave levam à conclusão de que o valor, por si só não é o mais importante. Isto é, a relação preço-qualidade (*good value*) influencia positivamente a valência. Por outro lado, as críticas negativas nesta categoria resumem-se às palavras-chave *expensive* (caro), *average* (mediocre) e *bad* (mau), o que atribui ênfase à importância da relação entre a qualidade e o preço

Observando a categoria do clima, denota-se que é um atributo pouco frequente nas críticas (0.5%), com uma valência muito positiva (0.79). As palavras-chave refletem uma das razões pelas quais o Algarve é popular no verão, o facto do clima ser agradável, estas palavras são *sunny* (soalheiro), *warm* (calor) e *hot* (quente).

A categoria “infraestruturas e atmosfera” inclui a ordenação do território, os acessos, disponibilidade de serviços, e a atmosfera sentida no local. As palavras-chave indicam a preferência por lugares calmos e pacíficos (*quiet*) e, pelo contrário, o excesso de pessoas (*crowded*) influencia negativamente a experiência dos turistas. Adicionalmente, observa-se um elogio à disponibilidade de serviços e/ou acessos (*plenty*), e a facilidade com que é possível aceder-lhes (*easy*). É um atributo evocado frequentemente nas críticas (9.7% do total de críticas) e que deixa boa impressão na mente do consumidor devido à valência de 0.79.

Curiosamente, a categoria “prestadores de serviços” mostrou-se ser o atributo mais evocado, tendo 19.5% dentro do total de críticas. Esta percentagem implica que os responsáveis pelos vários serviços disponibilizados no Algarve estão no centro da experiência do turista. No entanto, apesar da valência (0.76) implicar um resultado muito positivo, está entre as categorias com pontuação mais baixa, o que significa que há espaço para melhor. Não obstante, o rácio de críticas positivas está em linha com as restantes categorias (95%). As palavras-chave negativa *busy* (ocupado), *cold* (frio, - apatia) e *rude* indicam que os consumidores necessitam de um serviço atencioso, dedicado (corroborado pela palavra-chave positiva *helpful*) e amigável (*friendly*).

A última categoria, que engloba os residentes locais, é das menos frequentes (0.8%). A partir do rácio de comentários positivos (94%) e da valência (0.79) é possível perceber

que os turistas apreciam a presença de residentes locais. As palavras-chaves que definem este atributo são *lovely* (carinhoso) e *friendly* (amigável).

4.2. Imagem Afetiva

A imagem do destino não se restringe à dimensão cognitiva, é também composta pela dimensão afetiva (Gartner, 1994), que compreende os sentimentos despoletados pelos atributos presentes no destino (Agapito et al., 2013; Echtner e Ritchie, 1991; Hernández-Lobato et al., 2006). Como se pode verificar na tabela 4.1, para cada atributo da imagem é possível encontrar termos (maioritariamente adjetivos) que expressam um valor sentimental relativamente à experiência vivida. Sendo que um dos objetivos do presente estudo é compreender a imagem do destino no caso do Algarve, as críticas negativas e positivas foram separadas em grupos distintos, e a partir da frequência de palavras e da técnica de etiquetagem POS (*part-of-speech*) foi possível extrair somente os adjetivos mais frequentes nos dois grupos, juntamente com a respetiva frequência relativamente ao número total de palavras em toda a amostra e a média de valência entre as críticas que empenham o termo para descrever a experiência. Assim, a tabela 4.2 apresenta uma lista de vinte termos mais relevantes nas críticas positivas, e a tabela 4.3 apresenta os vinte termos mais relevantes nas críticas negativas. Ao extrair estes adjetivos pretende-se observar quais os termos mais comuns e, por conseguinte, compreender qual é a imagem afetiva do Algarve-

Tabela 4.2.

Vinte termos afetivos mais relevantes nas críticas positivas.

Termo	Frequência %	Média de valência	Termo	Frequência %	Média de valência
good	0.50	0.887	worth	0.17	0.884
great	0.48	0.928	perfect	0.16	0.923
nice	0.39	0.886	hot	0.14	0.878
excellent	0.27	0.917	amazing	0.11	0.934
beautiful	0.26	0.917	pleasant	0.11	0.930
friendly	0.23	0.930	safe	0.11	0.906
fantastic	0.20	0.921	enjoyable	0.11	0.947
wonderful	0.19	0.930	delicious	0.11	0.938
comfortable	0.19	0.925	fabulous	0.10	0.941
quiet	0.18	0.900	interesting	0.10	0.936

Os termos levantados na tabela 4.2 projetam os sentimentos positivos relacionados com a experiência turística no Algarve, a partir da mesma é possível observar os sentimentos positivos que surgem dos atributos do destino. A maioria dos termos está relacionado com a forma como se sente a atmosfera nos locais (*i.e., beautiful, fantastic, wonderful, comfortable, quiet, pleasant, safe e enjoyable*), enquanto outros termos realçam a qualidade dos atributos (*i.e., good, great, nice, excellent, perfect, amazing, delicious, fabulous e interesting*). O termo *hot* reflete, novamente, o clima no Algarve, e o termo *worth*, que no contexto de afetos estará relacionado com uma experiência ou produto valioso, demonstra a importância dada à relação preço-qualidade dos produtos e serviços turísticos. A valência dos termos serve para demonstrar que, de facto, estes termos têm uma influência importante na satisfação do consumidor.

Relativamente à tabela 4.3, que inclui os termos afetivos utilizados para descrever uma experiência negativa, encontra-se, também, termos relacionados com a avaliação da atmosfera no local (*i.e., uncomfortable, busy, loud e noisy*). No entanto, a maioria dos termos reflete a apreciação negativa de atributos (*i.e., bad, poor, cold, awful, cheap, average, expensive, worst, horrible e terrible*). Observam-se, também, termos que estão relacionados com a condição em que os locais e/ou produtos se encontram, como *old, basic, dirty e hard*. Por fim, o termo *hot* também se encontra entre os termos mais frequentes nas críticas negativas, o que demonstra a subjetividade do julgamento afetivo (Bigné et al., 2001; Echtner e Ritchie, 1991).

Tabela 4.3.

Vinte termos afetivos mais relevantes nas críticas negativas.

Termo	Frequência %	Média de valência	Termo	Frequência %	Média de valência
bad	0.30	-0.577	average	0.14	-0.535
old	0.28	-0.749	expensive	0.14	-0.573
poor	0.25	-0.692	worst	0.14	-0.849
basic	0.23	-0.444	dirty	0.12	-0.708
cold	0.19	-0.558	disappointed	0.12	-0.573
awful	0.17	-0.811	horrible	0.12	-0.886
cheap	0.16	-0.659	hot	0.12	-0.591
uncomfortable	0.15	-0.462	hard	0.12	-0.525
busy	0.15	-0.614	terrible	0.12	-0.824
loud	0.14	-0.729	noisy	0.11	-0.770

os turistas, assim como o termo *castle* demonstra uma atração construída pelo homem com interesse entre consumidores.

Apesar do alojamento ser uma categoria representada em apenas 6.4% das críticas (cf. Tabela 4.1), vários termos relacionados ao alojamento surgem dentro dos mais frequentes devido ao facto de que é comum ser referido pelos turistas em críticas a outros locais e serviços no Algarve. Isto é, termos como *room* podem estar relacionados com o alojamento, no entanto, é mais comum ser empenhado para descrever a dimensão dos espaços. Por exemplo:

The locals obviously love their beach, and it is busy but not overcrowded, always room to walk for miles.

Tabela 4.4. Cinquenta termos mais frequentes.
Adaptado de Marina-Roig e Huertas (2020: 6)

Termo	Frequência %	Termo	Frequência %
beach	0.66	excellent*	0.29
hotel	0.66	sea	0.29
room	0.61	recommend**	0.29
tour	0.57	local	0.29
good*	0.50	boat	0.29
great*	0.48	hour	0.28
restaurant	0.43	see	0.28
nice*	0.40	small	0.27
food	0.40	beautiful*	0.27
place	0.39	minute	0.26
trip	0.38	visit	0.26
clean	0.37	best	0.25
staff	0.36	service	0.25
town	0.36	experience	0.25
lovely*	0.36	helpful	0.24
view	0.35	old	0.24
walk	0.35	amazing*	0.24
time	0.33	pool	0.23
night	0.33	parking	0.23
area	0.33	price	0.23
friendly	0.32	right	0.22
water	0.31	plenty	0.22
location	0.31	close	0.22
bar	0.30	worth	0.22
guide	0.30	lot	0.22

Nota. *: termos afetivos; **: termos conativos.

As críticas que mencionam recomendação foram filtradas e analisadas quanto à sua frequência relativamente a todas as críticas, e a média de valência nas críticas em que surgem. Esta informação é ilustrada na tabela 4.5.

Tabela 4.5.

Termos que mencionam recomendação positiva e negativa.

Recommend	Frequência %	Valência
Positivo	12.57%	0.89
Negativo	0.73%	0.41

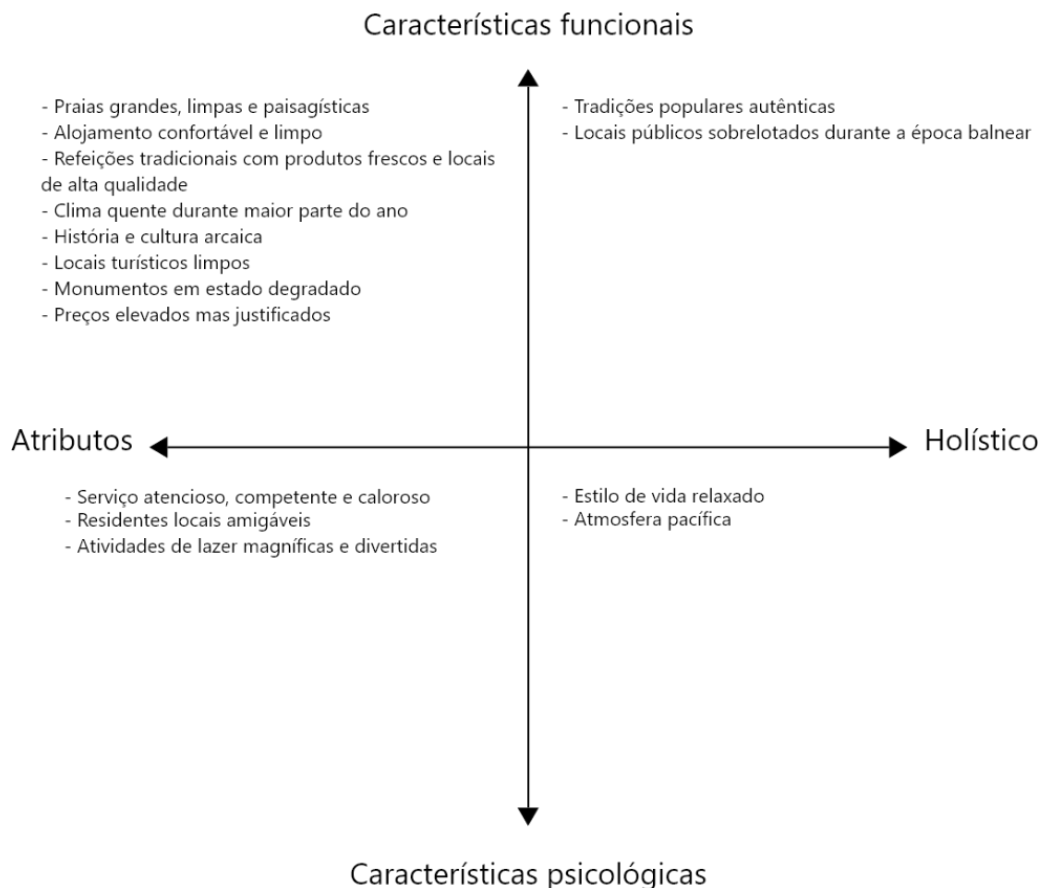
Na tabela 4.5 é possível verificar a influência da dimensão afetiva no comportamento futuro do consumidor (dimensão conativa). Isto é, a média da valência das críticas com recomendação positiva é de 0.89, enquanto a média da valência das críticas negativas é menor (0.41). Curiosamente, tendo em conta que o resultado da valência das críticas negativas é positivo, a recomendação negativa não implica que a experiência seja totalmente desagradável, mas que os sentimentos produzidos pela experiência não estão de acordo com a expectativa, resultando num desagrado e desaconselhamento.

4.4. Imagem global (holística)

Em termos globais (imagem holística), o Algarve é prezado pelos turistas devido às praias de grande dimensão, devidamente cuidadas; o alojamento adequado; refeições tradicionais compostas por produtos de alta qualidade; o clima agradável, que durante a época balnear é muito quente, e mesmo durante a época baixa encontram-se testemunhos que comentam positivamente a temperatura climática; apesar de haver pouca exposição à história e cultura portuguesa no Algarve, os turistas têm noção de que remete para centenas de anos atrás na história da humanidade, o que provoca admiração e interesse; encontram-se, também, vários testemunhos que criticam a falta de investimento em manter os monumentos históricos em boas condições; a higiene é um fator importante para a experiência positiva dos turistas, e no Algarve há unanimidade na opinião de que os espaços estão devidamente limpos; apesar de o poder de compra dos estrangeiros ser,

geralmente, superior à média portuguesa, os turistas consideram que os preços praticados no Algarve são altos, no entanto justificados na qualidade dos produtos; a prestação do serviço é feita de um modo atencioso, competente e animado; as atividades de lazer levam os turistas a sensações de magnificência e diversão; os modos e tradições populares são pouco expostos aos turistas, não obstante, são vistas como experiências autênticas; os turistas procuram experienciar um estilo de vida relaxante devido à atmosfera pacífica do Algarve, - em muitas ocasiões é empregue adjetivos como *quiet* (calmo), *pleasant* (agradável) e *comfortable* (confortável). Baseado no modelo proposto por Echtner e Ritchie (1991), a imagem global do destino ilustra-se na figura 4.2:

Figura 4.2. Imagem holística do Algarve



Fonte: Adaptado de Echtner e Ritchie (1991: 43)

5. DISCUSSÃO

O principal objetivo do presente estudo é a descrição da imagem do destino percebida pelos turistas internacionais no Algarve segundo o constructo tridimensional proposto na literatura (Agapito et al., 2013; Beerli e Martín, 2004; Bigné et al., 2001; Gartner, 1994; Jiang et al., 2021; Marine-Roig e Huertas, 2020), que compreende a imagem em três componentes (cognitivo-afetivo-conativo) que estabelecem uma relação hierárquica de influência.

Efetivamente, a utilização de técnicas de processamento de linguagem natural permitiu, sucessivamente, compreender a imagem do destino nestas três dimensões.

O presente estudo demonstra que os turistas tendem a referir mais frequentemente atributos cognitivos da imagem do destino do que as restantes dimensões do constructo, algo que é ilustrado a partir da frequência de palavras (*cf.* Tabela 4.4). Isto é, os turistas internacionais comentam, com regularidade, atributos do Algarve relacionados com as atrações naturais, alojamento, restauração, gastronomia e prestação dos serviços. Esta descoberta é semelhante ao que consta em Jiang et al. (2021), que comenta que, relativamente à imagem global de Hong Kong, observa-se uma grande concentração da imagem nas características e atributos funcionais. Semelhantemente, em Marine-Roig e Huertas (2020) verifica-se, também, uma maior presença e frequência de atributos relativos à imagem cognitiva do destino.

Em Hernández-Mogollón et al. (2014) é discutido que as motivações do turista face aos destinos podem ser de ordem física (entretenimento, bem-estar e relaxamento), cultural (outras culturas ou tradições), interpessoal (conhecer novas pessoas ou reencontrar amigos/família) e social (prestígio). Foi possível observar dados que demonstram que os turistas exprimem opiniões de (1) ordem física e (2) cultural:

1. O lazer e entretenimento é das categorias mais frequentes na imagem cognitiva e encontram-se termos afetivos relacionados com o bem-estar e relaxamento (*comfortable, quiet, pleasant, enjoyable*). Ilustrando, assim, que os turistas sentem motivações de ordem física em relação ao Algarve.

2. A frequência combinada do conjunto de categorias das atrações humanas, a história e cultura de Portugal e os costumes populares representa perto de um décimo de todas as críticas. São, precisamente, as categorias com um rácio de críticas positivas mais alto. Tornando, assim, as motivações de ordem cultural um fator presente na imagem do Algarve.

No entanto, as motivações de ordem social e interpessoal não foram possíveis de identificar no caso da imagem do Algarve.

Os resultados obtidos a partir da frequência de categorias referentes a componentes/atributos da imagem do destino (*cf.* Tabela 4.1) nas críticas turísticas permitem ilustrar que, de facto, a imagem pode ser, inicialmente, compreendida pelos atributos cognitivos do destino. A partir desta listagem possibilita-se a posterior compreensão da dimensão afetiva da imagem, que é, especificamente, ilustrada a partir de palavras, tipicamente adjetivos (Marine-Roig e Huertas, 2020), que expressam uma apreciação (positiva ou negativa). Estas palavras foram possíveis de obter a partir da junção da técnica de pré-processamento POS (*Part-of-speech Tagging*) e a frequência obtida pelo método TF-IDF (*term-frequency Inverse Document Frequency*), assim como é elaborado em Marine-Roig e Huertas (2020).

Com efeito, os resultados demonstram que os afetos sentidos perante determinada característica cognitiva expressam um julgamento afetivo (Baloglu e Love, 2005), o que indica que sem a imagem cognitiva não se pode observar a imagem afetiva, suportando a teoria de Gartner (1994) quanto à relação hierárquica do constructo. Isto é, os dados suportam o que é estipulado na literatura quanto à operação de uma avaliação dos atributos cognitivos segundo os sentimentos em relação aos mesmos (Agapito et al., 2013; Hernández-Lobato et al., 2006). Como se pode observar na Tabela 4.1, para cada categoria da imagem do destino atribui-se uma palavra-chave que, no fundo, implica uma avaliação em relação à categoria cognitiva. Esta estratégia é executada em Jiang et al. (2021), e no presente estudo demonstra-se que é um método aplicável ao conteúdo criado por turistas internacionais no Algarve.

A dimensão conativa foi captada a partir da discriminação de termos relacionados com recomendação em críticas positivas e negativas, o que permitiu obter resultados que

demonstram a influência da imagem afetiva na dimensão conativa (*i.e.*, comportamento do consumidor). Assim, observou-se que uma maior valência sentimental (expressão positiva) tem impacto na frequência com que a recomendação é efetuada. Segundo Woosnam et al. (2020), o comportamento do consumidor é resultado da realização das expectativas provocadas pela imagem cognitiva e afetiva do destino, algo que está espelhado nos resultados, - apesar de, no geral, a valência das críticas negativas que incluem recomendação ter um valor médio positivo, não se observa a recomendação com tanta frequência. Ou seja, não se trata de uma experiência totalmente negativa, mas de uma experiência que não cumpre as expectativas e desejos criados pela imagem do destino (Cardoso et al., 2019). Naturalmente, estas expectativas são subjetivas ao turista (Bigné et al., 2001; Echtner e Ritchie, 1991) e criadas durante o processo de seleção do destino.

Em Echtner e Ritchie (1991) é proposto um modelo que permite medir a imagem do destino em diferentes eixos (*cf.* Figura 4.2), nomeadamente, em três diferentes:

- Eixo que separa as características psicológicas das funcionais;
- Eixo que distingue elementos relativamente a características psicológicas ou funcionais, e elementos relativos à dimensão holística da imagem;
- Eixo que distingue os elementos únicos (psicológicas-holístico) e os elementos comuns (atributos-funcionais) da imagem do destino.

Segundo Agapito et al. (2010), relativamente à cidade de Lagos enquanto destino turístico (zona que se insere na região Algarve, a perspetiva holística dos turistas face ao destino refere as praias, o clima, a história e cultura, a postura amigável dos residentes, a atmosfera pacífica e segura. Comparando estes resultados com o presente estudo, é possível encontrar semelhanças, nomeadamente, no que está relacionado às praias, ao clima, à história e cultura, aos residentes e a atmosfera pacífica. Adicionalmente, destaca-se que no presente estudo observa-se que a imagem holística percebida pelos turistas internacionais inclui a autenticidade das tradições populares no Algarve, a crítica ao estado degradado em que se apresenta algumas atrações feitas pelo homem e a sobrelotação durante a época alta.

Destaca-se, ainda, que o Algarve demonstra ter como atributos únicos o estilo de vida relaxado e a atmosfera holística. Segundo Nowacki e Niezgoda (2020), características

únicas na imagem do destino contribuem para marcar uma vantagem competitiva em relação aos restantes destinos. A maioria dos atributos identificados na imagem do Algarve são comuns, o que significa que, tanto quanto consta, as características únicas do Algarve não são suficientemente populares.

Em Papadimitriou et al. (2018) é referido que a imagem global exerce influência na prática de *word-of-mouth* positivo. No caso do Algarve a frequência, valência e rácio de críticas positivas demonstram que as opiniões positivas no *TripAdvisor* ocupam a grande maioria. O que significa que a imagem global do destino algarvio é positiva, o que proporciona uma maior probabilidade de um turista demonstrar comportamento de lealdade perante o destino a partir de práticas como *word-of-mouth* positivo.

5.1. Implicações práticas para a gestão do setor do Turismo

A imagem do Algarve enquanto destino demonstra restrições presentes em algumas áreas do setor turístico. O presente estudo demonstra a importância, para as empresas, de realizar estudos críticos ao conteúdo gerado por turistas que tomaram a liberdade de aproveitar os produtos turísticos oferecidos no Algarve, devido ao facto de que o *word-of-mouth* representa um agente de influência orgânico, e é dos mais importantes para a tomada de decisão do consumidor. Tendo em conta de que a expectativa é um fator importante na satisfação do consumidor, e por conseguinte, no desempenho turístico (Afshardoost e Eshaghi, 2020; Hosany e Witham, 2010), é necessário compreender aquilo que é expresso pelo *word-of-mouth* para encontrar uma estratégia promocional que aproxime a expectativa à experiência. As empresas inseridas neste sector podem fazer uso destes métodos de estudos para avaliar continuamente a opinião do consumidor e tornar os respetivos esforços de marketing mais eficientes e, por conseguinte, com melhores resultados.

As empresas dentro deste setor e câmaras municipais, cujo produto que oferecem está relacionado com história e cultura, tradições populares e atrações contruídas pelo homem (e.g. museus e monumentos) devem aproveitar o facto de que são produtos altamente apreciados pelos turistas, e que, no entanto, encontram-se com pouca oferta no mercado turístico. Isto é, as atrações construídas pelo homem encontram-se em estado degradado, o que significa que se deve investir no melhoramento das condições em que os mesmos

estão apresentados; a história e cultura portuguesa presente no Algarve é algo que suscita interesse nos turistas e, no entanto, existe uma fraca oferta deste tipo de produto turístico; existe um número limitado de serviços turísticos que aproveite o impacto que as tradições e costumes populares têm nos turistas.

Os prestadores de serviços representam a categoria mais referida pelos turistas, o que demonstra a importância que este fator representa em qualquer serviço turístico. Ou seja, de modo a obter uma maior satisfação e imagem do Algarve, as empresas devem ter atenção, investir e melhorar continuamente a qualidade do serviço prestado ao cliente para estimular a competitividade do setor.

5.2. Contributo para a comunidade científica

O presente estudo contribui para a literatura científica disponível no tema da Imagem do Destino ao testar empiricamente se os modelos conceptuais tradicionais propostos até à data são eficientes na compreensão da imagem percebida pelos turistas, nomeadamente, o teorema tridimensional de Gartner (1994) e o modelo de medida da imagem do destino de Echtner e Ritchie (1991). Adicionalmente, contribui-se para a literatura académica demonstrando que o desenvolvimento teórico acerca da imagem do destino mantém-se relevante para os problemas e obstáculos dos dias atuais.

Os novos meios de comunicação permitem aos consumidores partilhar e obter opiniões acerca de determinado local turístico, o que afasta o controlo da promoção das entidades que participam no mercado. Estes meios, os media sociais, apresentam obstáculos para o estudo científico, como a grande dimensão dos dados e a presença de dados insignificantes, algo que é possível de superar a partir da combinação de técnicas de processamento de linguagem natural e algoritmos de aprendizagem supervisionada, assim como foi realizado no presente estudo para obter resultados significantes.

A metodologia adotada pelo autor demonstrou-se eficiente no tratamento de dados de grande dimensão e na produção de dados interpretáveis e significativos, e, apesar de ter sido utilizado em estudos passados para outros fins, esta investigação é a primeira a utilizá-lo para explorar a imagem do Algarve enquanto destino.

O presente estudo adiciona uma mais-valia à comunidade científica demonstrando como aplicar estas técnicas ao conteúdo textual não-estruturado, criado por internautas, para explorar a imagem do destino.

6. CONCLUSÃO

Ao longo desta dissertação pretendeu-se estudar e compreender a imagem do destino no caso do Algarve, com o objetivo de adotar uma metodologia que aproveita a tecnologia para lidar com um vasto número de dados. Esta metodologia foi, especificamente, a utilização de um algoritmo de análise de sentimentos e de técnicas de pré-processamento textual.

Durante o presente estudo foi apresentado várias descrições teóricas diferentes acerca da imagem do destino. Este conceito é, por natureza, imensamente complexo, devido ao facto dos destinos, enquanto produtos turísticos, serem um conjunto de características tangíveis e características não-tangíveis. O significado do conceito, devido à extensa pesquisa no tema, foi-se alterando ao longo das décadas.

Inicialmente, os estudos focavam-se apenas em fatores físicos e tangíveis dos locais turísticos, até que se demonstrou que fatores subjetivos ao indivíduo, como a personalidade e motivações também influenciavam o decorrer da experiência. A teoria principal que atingiu maior consenso propõe que a imagem do destino é formada a partir da interação entre três dimensões diferentes, que surgem segundo uma hierarquia, exercendo influência mútua. Cada dimensão enquadra-se numa etapa diferente na construção da imagem na mente do turista, tornando-se num ideal subjetivo a cada indivíduo.

Em suma, a imagem do destino é um constructo composto por três dimensões que partilham uma relação hierárquica de causa-efeito. Foi possível averiguar que, de facto, esta interação se verifica na experiência turística. A dimensão cognitiva reflete-se no conjunto de características funcionais, tangíveis e mensuráveis dos destinos, como, por exemplo, o valor monetário dos serviços, a gastronomia, a disponibilidade de acessos nos locais, as atrações turísticas e as atividades de lazer. De seguida, a dimensão afetiva comporta-se como um intermédio entre a imagem cognitiva e a imagem conativa, devido ao facto de que os turistas atribuem uma avaliação pessoal e subjetiva às características funcionais no destino, e esta avaliação terá impacto no comportamento futuro do mesmo perante o destino. Este comportamento despoletado pela experiência turística no destino designa-se de dimensão conativa, ou imagem conativa do destino, que surge como

resultado da interação entre o processamento cognitivo e afetivo dos estímulos presentes no local turístico, e pode ser observado segundo a intenção de recomendação, a repetição da visita turística e o comportamento de lealdade (*e.g. word-of-mouth*).

Como resultado do estudo, foi possível descobrir que o Algarve é prezado pelas suas atrações naturais, comida e restauração e atividades de lazer e entretenimento. Curiosamente, a categoria dos prestadores de serviço é o fator mais referido pelos turistas, o que significa que a alta qualidade do serviço nos diferentes produtos turísticos é uma condição imprescindível na satisfação do consumidor. Encontrou-se oportunidades no mercado turístico, como o melhoramento das atrações humanas (*i.e.* monumentos, museus, castelos e igrejas), e a necessidade de aumentar a exposição à história e cultura portuguesa por ser um fator muito apreciado pelos turistas. Nota-se, ainda, que o Algarve é visto como um destino em que é possível relaxar e viver um estilo de vida pacífico. Os turistas que visitam esta região exprimem regularmente que é um destino hospitaleiro, - os residentes locais e os profissionais dos serviços turísticos tornam a experiência de viagem numa vivência agradável e amigável.

Sucessivamente, foi possível explorar e compreender a imagem em três dimensões distintas que compõem o constructo da imagem do destino na mente do turista, - as dimensões cognitiva, afetiva e conativa.

A plataforma utilizada para obter os dados para esta dissertação, o *TripAdvisor*, demonstrou-se como uma fonte muito rica de testemunhos que permite obter informação que, tratada da forma correta, permite estudar detalhadamente a imagem do destino. Com efeito, a utilização de algoritmos de análise de sentimento produz bons resultados no que diz respeito à compreensão daquilo que é prezado e desprezado pelos turistas. No entanto, a valência obtida para cada opinião não é suficiente para tirar conclusões que englobem a grande escala dos dados. Para tal, é possível utilizar técnicas de processamento de linguagem natural para obter categorias que permitam ao investigador fazer sentido da opinião partilhada dos turistas.

6.1. Limitações e recomendações

O presente estudo ficou limitado aquilo que está presente na plataforma *TripAdvisor*. Ou seja, tendo em conta que os turistas partilham opiniões especificamente para determinado estabelecimento ou serviço, existe informação importante que pode não ser captada a partir desta plataforma. Nomeadamente, não foi possível compreender a opinião partilhada sobre a qualidade dos transportes na região do Algarve, nem descobrir as atrações únicas mais prezadas pelos turistas.

Para complementar e aprofundar a imagem do destino, é útil estudar e compreender aquilo que é comunicado, promovido e publicitado pelas entidades que participam no mercado turístico do Algarve. O presente estudo está limitado à imagem do destino percebida pelos turistas, e não inclui a imagem do destino projetada por estas instituições, algo que se recomenda para um futuro estudo sobre a imagem do Algarve enquanto destino.

A categorização dos dados foi decidida *à priori* e os dados foram classificados manualmente, o que pode tornar a categorização dos dados numa tarefa demasiado morosa no caso de ser necessário tratar de um número maior de instâncias de texto. Assim, recomenda-se que no futuro seja posto em prática uma abordagem diferente relativamente à segmentação textual, utilizando métodos de autoaprendizagem não supervisionada, como o modelo estatístico *Latent Dirichlet Allocation*.

O software utilizado para tratar os dados desta dissertação (*Orange Data Mining*) está limitado a modelos de classificação de uma única variável independente. Logo, é possível realizar um estudo em que as categorias sejam estipuladas *a priori*, e elaborado um *training set* que pode ser utilizado para treinar um modelo de classificação (*e.g.* regressão multivariada) que prevê quais as categorias que determinada instância textual pertence.

6.2. Trabalho futuro

Seria interessante complementar esta problemática com um estudo que, em vez de se centrar na opinião do consumidor, focar-se naquilo que é comunicado pelas entidades que participam no mercado do setor do turismo. Este tipo de comunicação pode estar presente

em blogues e redes sociais, e pode ser igualmente extraído pelo software utilizado no presente estudo (*WebHarvy*). Após a extração, é possível aplicar as categorias propostas na metodologia desta dissertação. Estar-se-ia a explorar a imagem projetada do Algarve enquanto destino turístico, e poder-se-ia comparar com os resultados presentes neste estudo. Complementado, assim, o estudo da imagem do destino no caso do Algarve com a compreensão da imagem projetada, podendo tirar mais conclusões relevantes para o tecido empresarial do setor turístico.

É possível utilizar a plataforma *TripAdvisor*, assim como foi feito neste estudo, para extrair opiniões de turistas nacionais e contrastar a diferença entre as duas imagens percebidas. Alternativamente, em Phelps (1986) é argumentada a possibilidade de que os residentes nos locais turísticos possuem uma imagem do local onde residem diferente dos turistas, e que esta está mais próxima da realidade. Seria interessante inquirir uma amostra de residentes de diferentes localidades da região do Algarve relativamente às categorias adotadas no presente estudo (atrações, atividades, ambiente e pessoas) e compreender o quão distante está a imagem percebida pelos turistas internacionais, da imagem presente na mente dos residentes locais.

Por outro lado, um estudo que, em vez de estipular a categorização do texto *a priori*, realize uma segmentação baseada em modelos de autoaprendizagem não-supervisionada, como o *Latent Dirichlet Allocation* para extrair tópicos dentro do conteúdo textual. Este método permite lidar uma amostra de dados muito mais volumosa, obter resultados interessantes, e explorar a possibilidade de utilizar métodos probabilísticos para segmentar dados textuais.

Tendo em conta que os resultados do presente estudo indicam que a categoria mais referida pelos turistas internacionais no Algarve é a dos prestadores de serviço, seria interessante inquirir uma amostra de turistas internacionais, especificamente acerca da qualidade do serviço que lhes foi prestado durante a estadia.

Por fim, baseado na descoberta de que as atrações históricas e culturais do Algarve poderão estar a ter pouca exposição e pouco interesse de investimento pelas entidades responsáveis, realizar um estudo que procure medir o interesse dos turistas internacionais

face a este tipo de atrações. Tirar conclusões sobre este tipo de produto turístico poderia indicar oportunidades e lacunas no turismo do Algarve.

REFERÊNCIAS BIBLIOGRÁFICAS

- Afshardoost, M., & Eshaghi, M. S. (2020). Destination image and tourist behavioural intentions: A meta-analysis. *Tourism Management*, 81(May), 104154. <https://doi.org/10.1016/j.tourman.2020.104154>
- Agapito, D., Oom do Valle, P., & da Costa Mendes, J. (2010). DESTINATION IMAGE Perspectives of Tourists versus Residents. *European Journal of Tourism, Hospitality and Recreation*, 1(1), 90–109.
- Agapito, D., Oom do Valle, P., & da Costa Mendes, J. (2013). The Cognitive-Affective-Conative Model of Destination Image: A Confirmatory Analysis. *Journal of Travel and Tourism Marketing*, 30(5), 471–481. <https://doi.org/10.1080/10548408.2013.803393>
- Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2014). Sentiment Analysis: Towards a Tool for Analysing Real-Time Students Feedback. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, 2014-Decem*(December), 419–423. <https://doi.org/10.1109/ICTAI.2014.70>
- Amersdorffer, D., Bauhuber, F., & Oellrich, J. (2012). The economic and cultural aspects of the social web: Implications for the tourism industry. *Journal of Vacation Marketing*, 18(3), 175–184. <https://doi.org/10.1177/1356766712449364>
- Baloglu, S., & Love, C. (2005). A cognitive-affective positioning analysis of convention cities: An extension of the circumplex model of affect. *Tourism Analysis*, 9(4), 299–308. <https://doi.org/10.3727/108354205789807210>
- Baloglu, S., & McCleary, K. W. (1999). U.S. international pleasure travelers' images of four Mediterranean destinations: A comparison of visitors and nonvisitors. *Journal of Travel Research*, 38(2), 144–152. <https://doi.org/10.1177/004728759903800207>
- Beerli, A., & Martín, J. D. (2004). Factors influencing destination image. *Annals of Tourism Research*, 31(3), 657–681. <https://doi.org/10.1016/j.annals.2004.01.010>
- Bigné Alcañiz, E., Sánchez García, I., & Sanz Blas, S. (2009). The functional-psychological continuum in the cognitive image of a destination: A confirmatory analysis. *Tourism Management*, 30(5), 715–723. <https://doi.org/10.1016/j.tourman.2008.10.020>
- Bigné, J. E., Sánchez, M. I., & Sánchez, J. (2001). Tourism image, evaluation variables and after purchase behaviour: Inter-relationship. *Tourism Management*, 22(6), 607–616. [https://doi.org/10.1016/S0261-5177\(01\)00035-8](https://doi.org/10.1016/S0261-5177(01)00035-8)
- Cai, L. A. (2002). Cooperative branding for rural destinations. *Annals of Tourism Research*, 29(3), 720–742. [https://doi.org/10.1016/S0160-7383\(01\)00080-9](https://doi.org/10.1016/S0160-7383(01)00080-9)
- Calantone, R. J., Di Benedetto, C. A., Hakam, A., & Bojanic, D. C. (1989). Multiple multinational tourism positioning using correspondence analysis. *Journal of Travel Research*, 28(2), 25–32. <https://doi.org/10.1177/004728758902800207>
- Cardoso, L., Dias, F., de Araújo, A. F., & Andrés Marques, M. I. (2019). A destination imagery processing model: Structural differences between dream and favourite destinations. *Annals of Tourism Research*, 74(November 2017), 81–94. <https://doi.org/10.1016/j.annals.2018.11.001>
- Carvalho, F. L., Sequeira, B. D., Wikesjö, M., & Ramos, C. M. Q. (2020). *The Unfolding*

- Theories on Destination Image, 1990-2020*. 313–338. <https://doi.org/10.4018/978-1-7998-3156-3.ch015>
- Chen, J. S. (2001). A case study of Korean outbound travelers' destination images by using correspondence analysis. *Tourism Management*, 22(4), 345–350. [https://doi.org/10.1016/S0261-5177\(00\)00066-2](https://doi.org/10.1016/S0261-5177(00)00066-2)
- Chon, K. (1992). The role of destination image in tourism: An extension. *The Tourist Review*, 47(1), 2–8. <https://doi.org/10.1108/eb058086>
- Crompton, J. L. (1979). Motivations for pleasure vacation. *Annals of Tourism Research*, 6(4), 408–424. [https://doi.org/10.1016/0160-7383\(79\)90004-5](https://doi.org/10.1016/0160-7383(79)90004-5)
- Dwivedi, M. (2009). Online destination image of India: A consumer based perspective. *International Journal of Contemporary Hospitality Management*, 21(2), 226–232. <https://doi.org/10.1108/09596110910935714>
- Dwivedi, M., Yadav, A., & Patel, V. R. (2009). The online destination image of Goa. *Worldwide Hospitality and Tourism Themes*, 1(1), 25–39. <https://doi.org/10.1108/17554210910949869>
- Echtner, C., & Ritchie, J. (1991). The meaning and measurement of destination image. *Journal of Tourism Studies*, 2(2), 2–12.
- El-Masri, M., Altrabsheh, N., & Mansour, H. (2017). Successes and challenges of Arabic sentiment analysis research: a literature review. *Social Network Analysis and Mining*, 7(1). <https://doi.org/10.1007/s13278-017-0474-x>
- Fakeye, P. C., & Crompton, J. L. (1991). Image Differences between Prospective, First-Time, and Repeat Visitors to the Lower Rio Grande Valley. *Journal of Travel Research*, 30(2), 10–16. <https://doi.org/10.1177/004728759103000202>
- Gallarza, M. G., Saura, I. G., & García, H. C. (2002). Towards a conceptual framework. *Management Fashion Adoption*, 29(1), 67–96. https://doi.org/10.1007/978-3-663-10707-1_3
- Gartner, W. C., & Hunt, J. D. (1987). An analysis of state image change over a twelve-year period (Utah 1971-1983). *Journal of Travel Research*, 26(2), 15–19. <https://doi.org/10.1177/004728758702600204>
- Gartner, William C. (1994). Journal of Travel & Image Formation Process. *Journal of Travel & Tourism Marketing*, 2(2–3), 191–216.
- Gedikoglu, Z. A., Backman, S. J., Mazer, J. P., & Backman, K. F. (2020). A Conceptual Framework to Understand Online Destination Images. 114–135. <https://doi.org/10.4018/978-1-7998-1947-9.ch008>
- Go, A., Bhayani, R., & Huang, L. (2009). Twitter Sentiment Classification using Distant Supervision. *Processing*, 1–6.
- Goldsmith, R. E. (2006). Eletronic Word-of-mouth. In *Encyclopedia of e-commerce, e-government and mobile commerce* (pp. 408–412). Idea Group Publishing.
- González-Rodríguez, M. R., Martínez-Torres, R., & Toral, S. (2016). Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. *International Journal of Contemporary Hospitality Management*, 28(11), 2609–2627. <https://doi.org/10.1108/IJCHM-02-2015-0057>
- Govers, R., & Go, F. M. (2008). Deconstructing Destination Image in the Information Age. *Information Technology & Tourism*, 6(1), 13–29.

- <https://doi.org/10.3727/109830503108751199>
- Graham M.S., D. (1996). Tourists' Images of a Destination-An Alternative Analysis. *Journal of Travel and Tourism Marketing*, 5(1–2), 41–45. <https://doi.org/10.1300/J073v05n01>
- Hernández-Lobato, L., Solis-Radilla, M. M., Moliner-Tena, M. A., & Sánchez-García, J. (2006). Tourism destination image, satisfaction and loyalty: A study in Ixtapa-Zihuatanejo, Mexico. *Tourism Geographies*, 8(4), 343–358. <https://doi.org/10.1080/14616680600922039>
- Hernández-Mogollón, J. M., Folgado-Fernández, J. A., & Duarte, P. (2014). Event Tourism Analysis and State of the Art. *European Journal of Tourism, Hospitality and Recreation*, 5(2), 83–102.
- Hosany, S., & Witham, M. (2010). Dimensions of cruisers' experiences, satisfaction, and intention to recommend. *Journal of Travel Research*, 49(3), 351–364. <https://doi.org/10.1177/0047287509346859>
- Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. *Proceedings of the National Conference on Artificial Intelligence*, 755–760.
- Huang, W., Zhu, S., & Yao, X. (2020). Destination Image Recognition And Emotion Analysis: Evidence From User-Generated Content Of Online Travel Communities. *The Computer Journal*, 00(00). <https://doi.org/10.1093/comjnl/bxaa064>
- Hunt, J. D. (1975). Image as a Factor in Tourism Development. *Journal of Travel Research*, 13(3), 1–7. <https://doi.org/10.1177/004728757501300301>
- Hunter, W. C. (2016). The social construction of tourism online destination image: A comparative semiotic analysis of the visual representation of Seoul. *Tourism Management*, 54, 221–229. <https://doi.org/10.1016/j.tourman.2015.11.012>
- Hutto, C.J. and Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for. *Eighth International AAAI Conference on Weblogs and Social Media*, 18. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109>
- Jiang, Q., Chan, C. S., Eichelberger, S., Ma, H., & Pikkemaat, B. (2021). Sentiment analysis of online destination image of Hong Kong held by mainland Chinese tourists. *Current Issues in Tourism*, 0(0), 1–22. <https://doi.org/10.1080/13683500.2021.1874312>
- Keller, K. L. (2013). *Strategic Brand Management: Building, Measuring, And Managing Brand Equity* (4th, Globa ed.). Pearson.
- Kim, S. E., Kim, H., & Yang, S. B. (2021). Why not this destination? The effects of travel constraints of independent travelers on destination image formation. *International Journal of Tourism Research*, June 2020, 1–13. <https://doi.org/10.1002/jtr.2463>
- Lalicic, L., Marine-Roig, E., Ferrer-Rosell, B., & Martin-Fuentes, E. (2021). Destination image analytics for tourism design: An approach through Airbnb reviews. *Annals of Tourism Research*, 86(November 2020), 103100. <https://doi.org/10.1016/j.annals.2020.103100>
- Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), 458–468. <https://doi.org/10.1016/j.tourman.2007.05.011>
- Liu, B. (2010). Sentiment analysis: A multifaceted problem. *IEEE Intelligent Systems*, 25(3), 76–80. <https://doi.org/10.1109/MIS.2010.75>

- Liu, M. T., Liu, Y., Mo, Z., & Ng, K. L. (2020). Using text mining to track changes in travel destination image: the case of Macau. *Asia Pacific Journal of Marketing and Logistics*, 33(2), 371–393. <https://doi.org/10.1108/APJML-08-2019-0477>
- Mak, A. H. N. (2017). Online destination image: Comparing national tourism organisation's and tourists' perspectives. *Tourism Management*, 60, 280–297. <https://doi.org/10.1016/j.tourman.2016.12.012>
- Marine-Roig, E., & Huertas, A. (2020). How safety affects destination image projected through online travel reviews. *Journal of Destination Marketing and Management*, 18(August). <https://doi.org/10.1016/j.jdmm.2020.100469>
- Mayo, E. (1975). Tourism and the National Parks: A Psychographic and Attitudinal Study. *Journal of Travel Research*, 14(1), 14–21. <https://doi.org/10.1177/004728757501400103>
- Nowacki, M., & Niezgodá, A. (2020). Identifying unique features of the image of selected cities based on reviews by TripAdvisor portal users. *Scandinavian Journal of Hospitality and Tourism*, 20(5), 503–519. <https://doi.org/10.1080/15022250.2020.1833362>
- Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31(1), 527–541. <https://doi.org/10.1016/j.chb.2013.05.024>
- Paik, J. H. (2013). A novel TF-IDF weighting scheme for effective ranking. *SIGIR 2013 - Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 343–352. <https://doi.org/10.1145/2484028.2484070>
- Pan, B., MacLaurin, T., & Crotts, J. C. (2007). Travel blogs and the implications for destination marketing. *Journal of Travel Research*, 46(1), 35–45. <https://doi.org/10.1177/0047287507302378>
- Pan, S., Lee, J., & Tsai, H. (2014). Travel photos: Motivations, image dimensions, and affective qualities of places. *Tourism Management*, 40, 59–69. <https://doi.org/10.1016/j.tourman.2013.05.007>
- Papadimitriou, D., Kaplanidou, K. (Kiki), & Apostolopoulou, A. (2018). Destination Image Components and Word-of-Mouth Intentions in Urban Tourism: A Multigroup Approach. *Journal of Hospitality and Tourism Research*, 42(4), 503–527. <https://doi.org/10.1177/1096348015584443>
- Park, E., Kang, J., Choi, D., & Han, J. (2020a). Understanding customers' hotel revisiting behaviour: a sentiment analysis of online feedback reviews. *Current Issues in Tourism*, 23(5), 605–611. <https://doi.org/10.1080/13683500.2018.1549025>
- Park, E., Kang, J., Choi, D., & Han, J. (2020b). Understanding customers' hotel revisiting behaviour: a sentiment analysis of online feedback reviews. *Current Issues in Tourism*, 23(5), 605–611. <https://doi.org/10.1080/13683500.2018.1549025>
- Peña, A. I. P., Jamilena, D. M. F., & Molina, M. Á. R. (2012). Validation of cognitive image dimensions for rural tourist destinations: A contribution to the management of rural tourist destinations. *Journal of Vacation Marketing*, 18(4), 261–273. <https://doi.org/10.1177/1356766712449351>
- Phelps, A. (1986). Holiday destination image - the problem of assessment. An example developed in Menorca. *Tourism Management*, 7(3), 168–180. [https://doi.org/10.1016/0261-5177\(86\)90003-8](https://doi.org/10.1016/0261-5177(86)90003-8)

- Ren, G., & Hong, T. (2017). Investigating online destination images using a topic-based sentiment analysis approach. *Sustainability (Switzerland)*, 9(10). <https://doi.org/10.3390/su9101765>
- Reza Jalilvand, M., Samiei, N., Dini, B., & Yaghoubi Manzari, P. (2012). Examining the structural relationships of electronic word of mouth, destination image, tourist attitude toward destination and travel intention: An integrated approach. *Journal of Destination Marketing and Management*, 1(1–2), 134–143. <https://doi.org/10.1016/j.jdmm.2012.10.001>
- San Martín, H., & Rodríguez del Bosque, I. A. (2008). Exploring the cognitive-affective nature of destination image and the role of psychological factors in its formation. *Tourism Management*, 29(2), 263–277. <https://doi.org/10.1016/j.tourman.2007.03.012>
- Stylidis, D., Shani, A., & Belhassen, Y. (2017). Testing an integrated destination image model across residents and tourists. *Tourism Management*, 58, 184–195. <https://doi.org/10.1016/j.tourman.2016.10.014>
- Stylos, N., Vassiliadis, C. A., Bellou, V., & Andronikidis, A. (2016). Destination images, holistic images and personal normative beliefs: Predictors of intention to revisit a destination. *Tourism Management*, 53, 40–60. <https://doi.org/10.1016/j.tourman.2015.09.006>
- Tasci, A. D. A., & Gartner, W. C. (2007). Destination image and its functional relationships. *Journal of Travel Research*, 45(4), 413–425. <https://doi.org/10.1177/0047287507299569>
- Um, S., & Crompton, J. L. (1992). The Roles of Perceived Inhibitors and Facilitators in Pleasure Travel Destination Decisions. *Journal of Travel Research*, 30(3), 18–25. <https://doi.org/10.1177/004728759203000303>
- Woosnam, K. M., Stylidis, D., & Ivkov, M. (2020). Explaining conative destination image through cognitive and affective destination image and emotional solidarity with residents. *Journal of Sustainable Tourism*, 28(6), 917–935. <https://doi.org/10.1080/09669582.2019.1708920>
- Yan, X., Wang, J., & Chau, M. (2015). Customer revisit intention to restaurants: Evidence from online reviews. *Information Systems Frontiers*, 17(3), 645–657. <https://doi.org/10.1007/s10796-013-9446-5>
- Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior*, 27(2), 634–639. <https://doi.org/10.1016/j.chb.2010.04.014>
- Yu, C. E., & Zhang, X. (2020). The embedded feelings in local gastronomy: a sentiment analysis of online reviews. *Journal of Hospitality and Tourism Technology*, 11(3), 461–478. <https://doi.org/10.1108/JHTT-02-2019-0028>
- Zhang, H., Fu, X., Cai, L. A., & Lu, L. (2014). Destination image and tourist loyalty: A meta-analysis. *Tourism Management*, 40, 213–223. <https://doi.org/10.1016/j.tourman.2013.06.006>

APÊNDICE A – *Workflow* aplicado utilizado no *Orange Data Mining*

