

A relação entre o ambiente urbano e o bem-estar: Análise em Lisboa, Portugal, utilizando redes sociais

The relation between urban environment and well-being: Analysis in Lisbon, Portugal, using social networks

Lúria Betco¹
Jorge Rocha²

Recebido 7 de julho de 2023; aceito 10 de novembro de 2023

RESUMO

Os problemas de saúde mental têm vindo a aumentar em todo o mundo, o que poderá estar associado ao crescimento da população urbana e ao estilo de vida a ela associado. O reconhecimento de que os diversos aspetos do ambiente urbano podem afetar a saúde mental dos indivíduos tem vindo a aumentar, uma vez que, estes são responsáveis por facilitar ou inibir comportamentos e estilos de vida que impactam o sentimento. Neste contexto é importante compreender o potencial impacto do ambiente urbano da cidade de Lisboa. Para tal recorreu-se à análise de sentimentos, utilizado um léxico do NRC Sentiment and Emotion, a partir de dados da rede social X, possibilitando a identificação dos locais em que tanto o sentimento positivo como negativo prevalecem. De seguida fez-se uso de um modelo de machine learning (ML) associado a um modelo-agnóstico de modo a aumentar a compreensão dos fatores do ambiente urbano que podem explicar o sentimento. Foram testados 4 modelos de ML, Random Forest (RF), Extreme Gradient Boosting (XGBoost), Neural Network (NN), o K-Nearest Neighbour (KNN) e um modelo linear para comparação (Generalized Linear Model-GLM). Os modelos agnósticos aplicados, o Local Interpretable Model-Agnostic Explanations (LIME) e o SHapley Additive exPlanation (SHAP), desempenharam um papel fundamental neste estudo. Respondendo à questão

¹ Centro de Estudos Geográficos; Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa, Portugal, e-mail: betcoiuria@edu.ulisboa.pt
ORCID: <https://orcid.org/0000-0001-8714-686X>

² Laboratório Associado Terra, Portugal, e-mail: jorge.rocha@edu.ulisboa.pt
ORCID: <https://orcid.org/0000-0002-7228-6330>

de partida, as variáveis explicativas que mais se relacionam com o sentimento são a distância a equipamentos fitness, a distância aos espaços verdes, a popularidade dos locais (estimada através da rede social Flickr) e a distância à rede ciclável.

Palavras-chave: bem-estar, redes sociais, análise de sentimento, morfologia urbana, machine learning.

ABSTRACT

Mental health problems have been rising worldwide, possibly associated with urban population growth and related lifestyles. The recognition that the various aspects of the urban environment can affect the mental health of individuals has been increasing since they are responsible for facilitating or inhibiting behaviors and lifestyles that impact the feeling. In this context, it is essential to understand the potential impact of the urban environment of the city of Lisbon. To do so, we resorted to sentiment analysis, using a lexicon from the NRC Sentiment and Emotion, based on data from the social network X, enabling the identification of places where both positive and negative sentiment prevail. Next, an machine learning (ML) model associated with an agnostic model was used to increase the understanding of the factors of the urban environment that can explain the sentiment. Four ML models were tested, Random Forest (RF), Extreme Gradient Boosting (XGBoost), Neural Network (NN), K-Nearest Neighbour (KNN), and a linear model for comparison (Generalized Linear Model - GLM). The agnostic models applied, the Local Interpretable Model-Agnostic Explanations (LIME) and the Shapley Additive exPlanation (SHAP), played a crucial role in this study. Answering the starting question, the explanatory variables most related to sentiment are distance to fitness facilities, distance to green spaces, the popularity of locations (estimated through the social network Flickr), and distance to the cycling network.

Key words: well-being; social networks; sentiment analysis; urban morphology; machine learning.

RESUMEN

Los problemas de salud mental han ido en aumento en todo el mundo, lo que puede estar relacionado con el crecimiento de la población urbana y el estilo de vida asociado a ella. Cada vez se reconoce más que los diversos aspectos del entorno urbano pueden afectar a la salud mental de las personas, ya que son responsables de facilitar o inhibir comportamientos y estilos de vida que repercuten en los sentimientos. En este contexto, es importante comprender el impacto potencial del entorno urbano de Lisboa. Para ello, se utilizó el análisis de sentimiento, utilizando un léxico del NRC Sentiment and Emotion, basado en datos de la red social X, lo que permitió identificar los lugares en los que prevalece tanto el sentimiento positivo como el negativo. A continuación, se combinó un modelo de aprendizaje automático (ML) con un modelo agnóstico con el fin de aumentar la comprensión de los factores del entorno urbano que pueden explicar el sentimiento. Se probaron cuatro modelos ML: Random Forest

(RF), Extreme Gradient Boosting (XGBoost), Neural Network (NN), K-Nearest Neighbour (KNN) y un modelo lineal de comparación (Generalized Linear Model - GLM). Los modelos agnósticos aplicados, el Local Interpretable Model-Agnostic Explanations (LIME) y el SHapley Additive exPlanation (SHAP), desempeñaron un papel fundamental en este estudio. En respuesta a la pregunta de partida, las variables explicativas más relacionadas con el sentimiento son la distancia a los gimnasios, la distancia a las zonas verdes, la popularidad de los lugares (estimada mediante la red social Flickr) y la distancia a la red ciclista.

Palabras clave: *bienestar, redes sociales, análisis de sentimientos, morfología urbana, aprendizaje automático.*

1. Introdução

Atualmente, mais de 50% da população em todo o mundo vive em áreas urbanas, é expectável que este número aumente para cerca de 70% em 2050 (United Nations, 2015). À medida que o mundo se torna cada vez mais urbanizado, planeadores e profissionais de saúde vão ficando preocupados com os impactes dos ambientes urbanos no bem-estar mental dos residentes. Apesar de a qualidade de vida em geral ter melhorado em todo o mundo, os problemas de saúde mental aumentaram, principalmente nos países industrialmente desenvolvidos (Hidaka, 2012), devido ao estilo de vida urbano (Reeves et al., 2011).

As evidências empíricas demonstraram que os fatores ambientais podem impactar positivamente e/ou negativamente o bem-estar mental, mas existem poucas diretrizes sobre como aumentar ou evitar certos ambientes (Hajrasoulih et al., 2018). Dado que a forma e a morfologia da cidade podem ser determinantes, é fundamental determinar quais os fatores que podem contribuir para alcançar o bem-estar, de modo a projetar a cidade nesse sentido. A questão adicional é, em que medida cada fator é quantitativamente influente, e qual desses critérios terá a maior prioridade e importância (Fathi et al., 2020).

A transformação do espaço urbano simultaneamente com a digitalização da sociedade, veio possibilitar o conhecimento acerca do quanto o bem-estar individual é influenciado pelo ambiente construído envolvente (Romice et al., 2016). Isto é possível, acedendo ao conteúdo gerado pelos internautas nas redes sociais, que fornece informações sobre os utilizadores e o seu ambiente, através de preferências, opiniões e interações (Aloufi et al., 2017).

As pessoas expressam diariamente o seu sentimento através de textos e imagens diretamente nas redes sociais (e.g., Facebook, Twitter e Flickr) (Wang et al., 2014), esta abundância de dados torna desafiante a análise e obtenção de insights, levando a um aumento da procura pela análise de sentimento (Afriliana et al., 2022).

2. Bem-estar

O conceito de bem-estar é ambíguo, carece de uma definição universalmente aceite e apresenta numerosas, e por vezes divergentes interpretações (McGillivray & Clarke, 2006). Pode ser visto como tendo duas dimensões:

objetivo e subjetivo. O bem-estar objetivo é uma medida de aproximação baseada nas necessidades e direitos humanos básicos, incluindo aspetos como a alimentação adequada, a saúde física, a educação e a segurança. O bem-estar subjetivo inclui aspetos como a satisfação com a vida, emoções positivas, e se a vida é significativa (Hajrasoulih *et al.*, 2018), é medido perguntando diretamente às pessoas como pensam e se sentem sobre o seu próprio bem-estar (Hajrasoulih *et al.*, 2018). Segundo Diener (1984). Uma pessoa com elevado bem-estar subjetivo apresenta níveis altos de satisfação com a vida, elevados sentimentos positivos e baixos sentimentos negativos (E. Diener, 1984).

O bem-estar subjetivo caracteriza-se por ser uma dimensão positiva da saúde. Sendo, simultaneamente, um conceito complexo (Galinha & Ribeiro, 2005), porque inclui três tipos de bem-estar interligados: a) hedónico (Deci & Ryan, 2008), b) psicológico (Huppert, 2009) e c) social (Keyes, 1998) (o bem-estar psicológico e social são dois aspetos do bem-estar eudemónico) (Hajrasoulih *et al.*, 2018). O bem-estar psicológico e social (eudemónico) relaciona-se com boas experiências para a pessoa, e tende a ser associado a efeitos a longo prazo e duradouros e o bem-estar hedónico é o mais primitivo e a curto prazo, resultante do prazer e da prevenção da dor (Hajrasoulih *et al.*, 2018).

2.1 Fatores determinantes para o bem-estar

O reconhecimento dos aspetos sociais, físicos e culturais, que podem afetar a saúde mental do individuo tem vindo a aumentar, uma vez que, estes são responsáveis por facilitar ou inibir comportamentos e estilos de vida que impactam o bem-estar (Santana *et al.*, 2007).

A forma urbana tem impacte nas três dimensões do bem-estar subjetivo, por exemplo, uma curta caminhada num parque local proporciona um prazer imediato (hedónico), transformar este comportamento numa rotina diária restaura a mente da fadiga mental (psicológica) e proporciona uma grande oportunidade de interação social com os vizinhos (social) (Hajrasoulih *et al.*, 2018). É possível sugerir estratégias de planeamento e design que promovam o bem-estar hedónico, psicológico e social através de intervenções na forma urbana (Hajrasoulih *et al.*, 2018).

A literatura revela a existência de lacunas significativas, ilustradas na Figura 1. Em alguns casos, estas lacunas resultam de informações contraditórias (e.g., a relação entre o modo de transporte e o bem-estar hedónico), enquanto noutros casos resulta de estudos que não conseguiram encontrar relações significativas (e.g., a relação entre o tipo de habitação e o bem-estar psicológico). Noutros casos, a publicação ou estudo não apresenta quaisquer dados (Hajrasoulih *et al.*, 2018).

O autor Hajrasoulih *et al.*, (2018) desenvolveu ainda um esquema conceptual, no qual apresenta o impacte da forma urbana no bem-estar mental. Tendo sido desenvolvido com o propósito de orientar os estudos de forma organizada e preencher as lacunas existentes. O esquema foi inspirado no quadro de sustentabilidade para o design urbano de Larco (2015), adaptado de forma a abordar a natureza complexa e multidimensional dos impactes no bem-estar mental (Hajrasoulih *et al.*, 2018). Szombathely *et al.* (2017) também

desenvolveram um modelo conceptual, no qual relacionam os determinantes de quatro sectores diferentes (individuo, sociedade, stressores e morfologia) e associam-nos ao bem-estar urbano relacionado com a saúde (UrbWellth) (Von Szombathely *et al.*, 2017).

		Bem-estar mental		
		Bem-estar hedónico	Bem-estar psicológico	Bem-estar social
Forma urbana	Vegetação	+ Ver e interagir com espaços verdes especialmente jardins ou áreas de paisagem	+ Proximidade a espaços verdes (todas as escalas)	+ Proximidade a espaços verdes >< Abundância de árvores
	Densidade	- Densidade superior	- Densidade superior	><
	Uso do solo		+ Acesso a instalações	+ Espaços comunitários, praças, espaços públicos Uso misto (evidência de conflito)
	Casa		><	+ Idade da habitação e diversidade
	Transporte	><	+ Curto trajeto de viagem	+ Andar mais a pé
	Configuração espacial	+ Vizinhança com design tradicional	+ Elementos fundamentais para o espaço publico abertos + Ambiente atrativo + Ambiente aberto + Arquitetura variada - Altura do edificado	- Colinas íngremes + Configuração semifechada + Mobiliário urbano
	Stressores	- Poluição do ar + Ruído >< Cores	- Proximidade às fontes de ruído (e.g., aeroporto) e indústrias perigosas	

Figura 1. Relação entre forma urbana e bem-estar mental; + correlação positiva, - correlação negativa, >< provas contraditórias (Fonte: Hajrasoulih *et al.*, 2018).

3. Área de estudo

Para este estudo seleccionou-se como área de estudo o município de Lisboa, que se subdivide em 24 freguesias representadas na Figura 2. A cidade de Lisboa,



Figura 2. Freguesias do município de Lisboa.

capital de Portugal, situa-se junto ao estuário do rio Tejo e apresenta uma área de 86,83 km². Em termos populacionais, alberga 509 515 habitantes (INE, 2021), apresentando uma densidade populacional de 5 868 habitantes por km². A ampla gama de atividades escolhidas pelo turista ao visitar Lisboa é um fator de vantagem competitiva, destacando-se a visita a monumentos e museus, os passeios a pé, a gastronomia e os vinhos, e ainda a vida noturna (Observatório Turismo de Lisboa, 2019).

4. Metodologia

Foi desenvolvido um modelo para compreender os múltiplos fatores que podem ajudar a explicar o sentimento, e como tal o bem-estar subjetivo hedónico em ambientes urbanos. Este modelo inclui quatro fases principais: 1) recolha e pré-tratamento de dados espaciais, 2) diagnóstico da multicolinearidade de dados, 3) construção do modelo ML e 4) aplicação de uma abordagem modelo-agnóstica para a interpretação. O processo de trabalho encontra-se ilustrado na Figura 3.

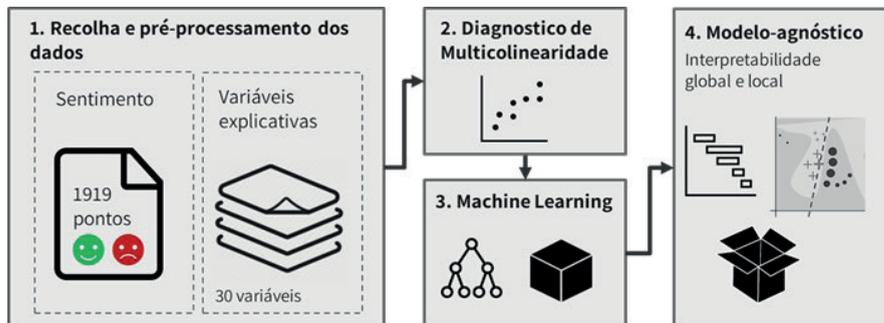


Figura 3. Metodologia do processo de modelação.

4.1 Recolha e pré-processamento dos dados

4.1.1 Variável dependente

Os dados do Twitter permitem uma abordagem mais económica e dinâmica do que o usual inquérito e, sobretudo, a capacidade de adicionar uma dimensão geométrica objetiva ao estudo subjetivo (Cao *et al.*, 2018). Além da informação espacial, o conteúdo dos tweets também fornece informações sobre as opiniões e sentimentos dos utilizadores, permitindo elaborar uma análise do sentimento (Balahr *et al.*, 2014). Portanto, o bem-estar subjetivo foi determinado a partir do sentimento expresso pelos utilizadores na rede social Twitter (atual X), que constituirá variável dependente no modelo ML preditivo.

4.1.1.1 Análise de sentimento

Esta fase da metodologia tem como objetivo identificar automaticamente os sentimentos expressos nos comentários publicados pelos utilizadores da rede social X, com recurso a um léxico de sentimentos, isto é, uma lista de palavras manualmente rotuladas com polaridade positiva e negativa. Que, deste modo, veio permitir calcular a pontuação geral do sentimento para cada publicação.

Optou-se pela utilização do léxico *NRC Word-Emotion Association Lexicon (EmoLex)*, porque dispõe de uma lista de palavras em inglês associadas a 8 emoções básicas (raiva, medo, antecipação, confiança, surpresa, tristeza, alegria e repulsa) e 2 sentimentos (negativo e positivo). Dispõe ainda de versões em mais de cem idiomas (Mohammad & Turney, 2015).

Embora, já existam trabalhos que se concentrem em palavras que ‘denotam’ emoção, o EmoLex desenvolvido por Saif Mohammad e Peter Turney (2013) impactou por incluir o maior conjunto de palavras associadas ou que conotam com uma emoção, sendo por isso o maior léxico deste tipo existente atualmente (Mohammad & Turney, 2015). A vantagem deste tipo de método é não precisar de dados de treino (como é o caso do método de machine-learning supervisionado) (Shayaa *et al.*, 2018).

Dos tuitos (*tweets*) publicados na cidade de Lisboa selecionaram-se apenas os georreferenciados e públicos, em que o utilizador pôde optar pelo fornecimento das coordenadas exatas. Assim, inicialmente dispunha-se de um conjunto de 16791 dados pontuais georreferenciados, correspondentes a comentários no Twitter publicados na cidade de Lisboa durante o ano de 2019. Destes selecionaram-se apenas os comentários publicados durante o dia, entre as 9h e 19h, ficando-se com 9446, dado ser as horas que correspondem a uma maior utilização do espaço urbano. Estes comentários foram posteriormente submetidos à análise de sentimento e classificação de emoções com recurso ao *EmoLex* no idioma português.

De modo a ser possível introduzir os dados referentes ao sentimento no modelo preditivo, foi necessário transformar a variável de resposta de um sistema de numeração natural para um sistema binário (0 ou 1). Como tal, classificou-se a pontuação resultante das palavras (portuguesas) associadas ao sentimento positivo (1) e negativo (0). Uma vez que a maioria dos comentários apresentava valores de sentimento neutros (*i.e.*, não apresenta pontuações de palavras portuguesas associadas a sentimentos), após implementada a condição ficou-se com apenas 1919 dados binários de sentimento.

4.1.2 Variáveis independentes

A seleção das 30 variáveis independentes (Tabela 1) foi efetuada a partir da literatura, do esquema conceptual da relação entre a forma urbana e bem-estar mental de Hajrasoulih *et al.*, (2018) e do modelo conceptual de bem-estar urbano relacionado com a saúde (UrbWellth) de Szombathely (2017) (Von Szombathely *et al.*, 2017), utilizando variáveis que se inserem nos setores: morfologia, stressores e sociedade.

Tabela 1. Variáveis independentes e a respetiva fonte dos dados

<i>Variável independente</i>	<i>Ano</i>	<i>Fonte dos dados</i>	<i>Bibliografia</i>
Popularidade dos locais (likes/km ²)	2019	Flickr	Marsillo <i>et al.</i> , 2022; J. Wang <i>et al.</i> , 2022
Riqueza relativa (Relative richness)	2018		
Diversidade (Diversity)	2018		
Dominância (Dominance Index)	2018	DGT, Carta de Uso e Ocupação do Solo	de la Riva <i>et al.</i> , 2023; Gai <i>et al.</i> , 2022; Marsillo <i>et al.</i> , 2022; Tao <i>et al.</i> , 2022; Zhang <i>et al.</i> , 2021
Fragmentação (Fragmentation index)	2018		
Radiação solar global	2017	Dados Lidar, CML	Jiao <i>et al.</i> , 2021
Largura das ruas	2017	Dados Lidar, CML	Gjerde & Vale, 2022; Mayor & Coleman, 2011
WorldView Built-Up Index (WV-BI)	2017		
WorldView Non-Homogeneous Feature Difference (WV-NHFD)	2017	WorldView-2	He <i>et al.</i> , 2023; Huang <i>et al.</i> , 2022; Schwartz <i>et al.</i> , 2019; Tao <i>et al.</i> , 2022; R. Wang <i>et al.</i> , 2020
WorldView Water Index (WV-WI)	2017		
NDVI	2017		
Poluente O3	2019		
Poluente NO2	2019	Agência Portuguesa do Ambiente (APA)	Du <i>et al.</i> , 2018; He <i>et al.</i> , 2023; R. Wang <i>et al.</i> , 2020; Xiong <i>et al.</i> , 2020
Poluente PM10	2019		
Ruído Global	2020	Câmara Municipal de Lisboa	Abdul-Rahman <i>et al.</i> , 2021; Clark <i>et al.</i> , 2020; Jiao <i>et al.</i> , 2021; Lawton & Fujiwara, 2016
Rede ciclável	2021	Câmara Municipal de Lisboa	Cebeci <i>et al.</i> , 2023; Lai & Deal, 2023; Leyland <i>et al.</i> , 2019
Desporto Equipamentos Fitness	2021	Câmara Municipal de Lisboa	Eddolls <i>et al.</i> , 2018
Corredor Verde	2021	Câmara Municipal de Lisboa	Chen <i>et al.</i> , 2022; Ghahramani <i>et al.</i> , 2021
Elementos de Água	2021	Câmara Municipal de Lisboa	Tao <i>et al.</i> , 2022; Völker & Kistemann, 2011; White <i>et al.</i> , 2020
Espaços Verdes	2021	Câmara Municipal de Lisboa	Ayala-Azcárraga <i>et al.</i> , 2019; Chen <i>et al.</i> , 2022; Ghahramani <i>et al.</i> , 2021

<i>Variável independente</i>	<i>Ano</i>	<i>Fonte dos dados</i>	<i>Bibliografia</i>
Imóveis e Monumentos de Interesse Público	2021		
Imóveis, Monumentos e Conjuntos de Interesse Municipal	2021	Câmara Municipal de Lisboa	Abdelmoula & Bouayed Abdelmoula, 2023; Aloufi <i>et al.</i> , 2017; Stellacci & Moro, 2022
Monumentos Nacionais	2021		
Grandes parques e Jardins de Lisboa	2021	Câmara Municipal de Lisboa	Lai & Deal, 2023; Schwartz <i>et al.</i> , 2019; Yuen & Jenkins, 2020
Centros Comerciais	2021	Câmara Municipal de Lisboa	Benabbou & Lee, 2019; Gao <i>et al.</i> , 2022
Centros de Saúde	2021	Câmara Municipal de Lisboa	Harrison <i>et al.</i> , 2012; Ye <i>et al.</i> , 2022
Alojamentos Hoteleiros	2021	Câmara Municipal de Lisboa	Lipovčan <i>et al.</i> , 2014
Alojamentos Locais	2019	Sistema de Informação Geográfica do Turismo (SIGTUR)	Lipovčan <i>et al.</i> , 2014; Martín-Martín <i>et al.</i> , 2023
Criminalidade	2019	Polícia de Segurança Pública de Lisboa	Lorenc <i>et al.</i> , 2012; Ristea <i>et al.</i> , 2021; Stellacci & Moro, 2022
Idade média ponderada do edificado	2011	BGRI, Instituto Nacional de Estatística (INE)	Abdelmoula & Bouayed Abdelmoula, 2023; Reilly <i>et al.</i> , 2018

4.2 Diagnóstico de multicolinearidade

A segunda fase da abordagem do modelo abrangeu uma análise estatística preliminar, durante a qual se diagnosticou a multicolinearidade das variáveis calculando o fator de inflação da variância (VIF) (Dohoo *et al.*, 1997; Lin, 2008). Calculou-se o VIF utilizando a biblioteca CARET (Classification and Regression Training) do software estatístico R, permitindo assim excluir as variáveis altamente correlacionadas (Naimi *et al.*, 2014). Qualquer variável com um VIF > 5 foi excluída do modelo (G. James *et al.*, 2013; Johnston *et al.*, 2018). Após o cálculo do VIF e a exclusão das variáveis altamente correlacionadas o número final de variáveis incluídas no modelo foi de 10 (exemplificadas na Figura 4).

4.3 Construção do modelo de Machine Learning

O Machine Learning (ML), que é um subcampo da inteligência artificial (IA), supera com sucesso as limitações dos métodos estatísticos. Em comparação aos métodos tradicionais, o ML é reconhecido por alcançar exatidões superiores ou pelo menos equivalentes (Lima *et al.*, 2015; Ren *et al.*, 2020; Shortridge *et al.*, 2016). Por sua vez, as abordagens ML têm muitas vantagens, como a capacidade de lidar com dados de diferentes tipos, estruturas e quantidades (*i.e.*, big data) (Molnar, 2019), não sendo sensíveis à escala das variáveis (o que significa que não há necessidade de as normalizar) (Viana *et al.*, 2021); por conseguinte, é possível combinar dados de diversas fontes para modelar relações não lineares complexas que descrevam a variação do sentimento em ambientes urbanos.

A terceira fase envolveu a construção do modelo de predição utilizando os seguintes algoritmos: Random Forest (RF), Generalized Linear Models (GLM), Neural Networks (NNET), k-Nearest Neighbors (KNN) e o XGBoost (XGB). Os algoritmos foram executados utilizando novamente a biblioteca CARET. Optou-se então por uma divisão dos dados de 90% para o treino e 10% para o teste, o primeiro é responsável pela aprendizagem do modelo e o segundo pela realização dos ajustes.

Tabela 2. Desempenho dos algoritmos ML na predição do sentimento

Algoritmo	Exactidão	Sensibilidade	Especificidade	Predição de positivos	Predição de negativos
RF	0,8125	0,8095	0,8140	0,6800	0,8974
GLM	0,7188	0,2222	0,9612	0,7368	0,7168
NNET	0,7344	0,2222	0,9845	0,8750	0,7216
KNN	0,8021	0,8095	0,7984	0,6623	0,8957
XGB	0,8021	0,8254	0,7907	0,6582	0,9027

Avaliou-se o desempenho dos quatro modelos de machine learning e do modelo linear (GLM) na predição do sentimento. Dos cinco modelos representados na Tabela 2, o algoritmo de *machine learning* RF foi o que alcançou o maior desempenho com uma exatidão de 0,81. Para além da construção do modelo a partir do RF, criou-se ainda um outro modelo de ML utilizando o algoritmo XGB por este permitir aplicar o método de interpretação local SHAP e observar sobre a forma de gráfico a agregação dos *Shapley values* (Lingjærde *et al.*, 2021).

4.4 Modelos-agnósticos

Apesar das potenciais vantagens, os algoritmos ML permanecem principalmente sob uma formulação de 'caixa negra', o que significa que sem mais intervenções, não é possível interpretar ou reconstituir diretamente como um modelo executa a inferência ou predição devido aos muitos pesos internos ou informações estruturais (Molnar, 2019).

A Inteligência artificial explicável surgiu recentemente como uma importante área de investigação, que propõe medidas estatísticas avançadas e ferramentas de visualização para melhorar a interpretação do ML (Carvalho *et al.*, 2019; Molnar, 2019; Murdoch *et al.*, 2019). Este é um aspeto crítico quando os resultados do modelo ML são usados para a tomada de decisão (Ribeiro *et al.*, 2016).

No caso dos modelos complexos (e.g., RF e XGB) não podemos usar o modelo original como explicação porque não é fácil de entender. Em vez disso, utiliza-se um modelo de explicação mais simples, que definimos como uma aproximação interpretável do modelo original (Lundberg & Lee, 2017).

Assim, esta fase envolveu a interpretação do modelo ML utilizando a abordagem dos modelos-agnósticos (MA) para aferir a importância das variáveis explicativas. A abordagem MA fornece uma explicação com base nos diferentes comportamentos dos modelos complexos ajustados apresentando a informação a nível global (i.e., o que o modelo aprendeu com as variáveis de entrada) e a nível local (i.e., os raciais que o modelo fornece para cada estimativa) (Viana *et al.*, 2021).

Neste estudo, aplicaram-se modelos agnósticos de métodos globais e locais que facilitam a interpretação global do modelo e a interpretação de observações locais. Para interpretação global do modelo implementou-se os métodos PFI (Permutation Feature Importance) e PDP (Partial Dependence Plot) e para a interpretação local utilizou-se os métodos LIME (Local Interpretable Model-agnostic Explanations) e o SHAP (Shapley Additive ExPlanations) (Lundberg & Lee, 2017).

5. Análise e discussão dos resultados

5.1 Permutation Feature Importance (PFI)

Os outputs da modelação RF foram examinados utilizando o método PFI para a interpretação global (Figura 4). O método PFI permitiu compreender quais as variáveis que contribuíram para os resultados subjacentes ao modelo ML e quantificar a sua importância.

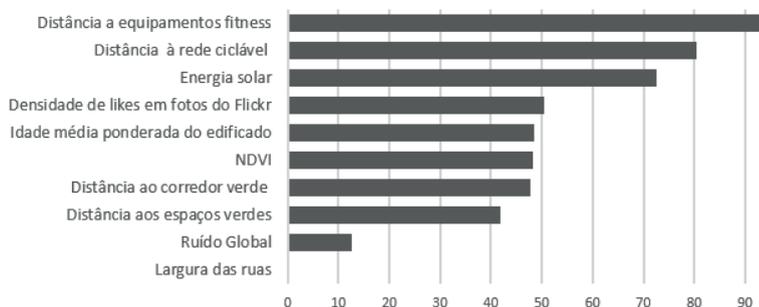


Figura 4. As 10 variáveis mais influentes na explicação do sentimento segundo o método global PFI.

A Figura 4, apresenta as 10 variáveis mais influentes na explicação do sentimento (com uma pontuação de importância > 50). Estas foram a distância a equipamentos fitness, distância à rede ciclável, energia solar e a popularidade dos locais (traduzida por uma grande concentração de likes em fotos do Flickr).

5.2 Partial Dependence Plot (PDP)

O método PDP foi também utilizado para a interpretação global da modelação RF. O PDP é um método de interpretação global, em que o gráfico apresenta o efeito marginal de uma única feature sobre o sentimento predito de um modelo anteriormente ajustado (Friedman, 2001).

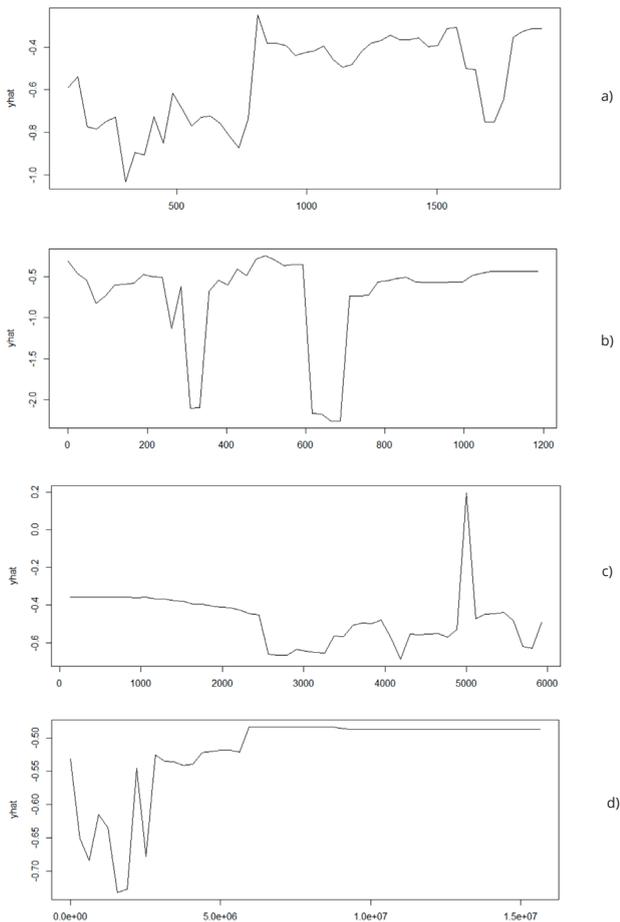


Figura 5. Curvas de resposta das primeiras 4 variáveis mais influentes no incremento do sentimento segundo o método global PDP: a) distância a equipamentos fitness; b) distância à rede ciclável; c) energia solar; d) popularidade dos locais.

A Figura 5 apresenta as curvas de resposta para as primeiras quatro variáveis mais influentes e o quanto estas se traduzem no incremento positivo ou negativo do sentimento. Analisando cada gráfico em particular, pode-se verificar que: (a) a proximidade aos equipamentos fitness não influencia positivamente o sentimento, verificando-se uma preferência por distâncias maiores a estes; (b) a proximidade à rede ciclável influencia de um modo geral positivamente o sentimento; (c) valores muito baixos de energia solar traduzem-se em valores moderados de sentimento, o pico nos 5000 kw-2 traduz uma preferência por espaços com maior exposição solar; (d) a popularidade dos locais (likes/km²) influencia positivamente o sentimento, verificando-se uma preferência por locais mais populares e/ou visualmente mais atrativos.

5.3 Local Interpretable Model-agnostic Explanations (LIME)

O método LIME foi utilizado para a interpretação local do modelo de caixa negra *Random Forest*, fornecendo uma explicação ao nível da observação (Elshawi *et al.*, 2019). Este método explica uma predição substituindo o modelo complexo por um modelo de substituição localmente interpretável (Molnar, 2019).

Efetuuou-se uma representação visual das primeiras duas observações individuais (7 e 23) dos dados do subconjunto local, por serem as que melhor justificam os resultados (Figura 6). Cada gráfico apresenta a probabilidade predita de cada observação como sentimento positivo ou negativo. Da mesma forma, apresenta as cinco variáveis que mais contribuem para a explicação da predição do modelo no local de cada observação e se estas variáveis aumentam (suportam) ou diminuem (contradizem) a probabilidade de sentimento (Salami *et al.*, 2020).

A Figura 6-a) apresenta a explicação da predição da observação 7 como sentimento positivo (*label* = 1), com uma forte probabilidade de 0,81. Com estes resultados, podemos deduzir que, para o caso 7 a probabilidade de ser um sentimento positivo é de 81%. Neste caso, a explicação LIME foi criada com base nas cinco variáveis: *Normalized Distance Vegetation Index* (NDVI), densidade de likes (popularidade do local), distância à rede ciclável, distância aos espaços verdes e distância aos equipamentos fitness. Estas contribuem positivamente para a explicação do sentimento positivo, à exceção do NDVI que contradiz a felicidade. Os valores de NDVI inferiores a 0,0352 referem-se a áreas com pouca ou nenhuma clorofila (*e.g.*, áreas artificializadas), estando negativamente correlacionadas com o sentimento positivo.

A tristeza (sentimento negativo; *label*= 0) é mais difícil de prever que a felicidade (sentimento positivo; *label*= 1), por apresentar valores de probabilidade inferiores (0,19 e 0,24) como se pode verificar na Figura 6-b e 6-d).

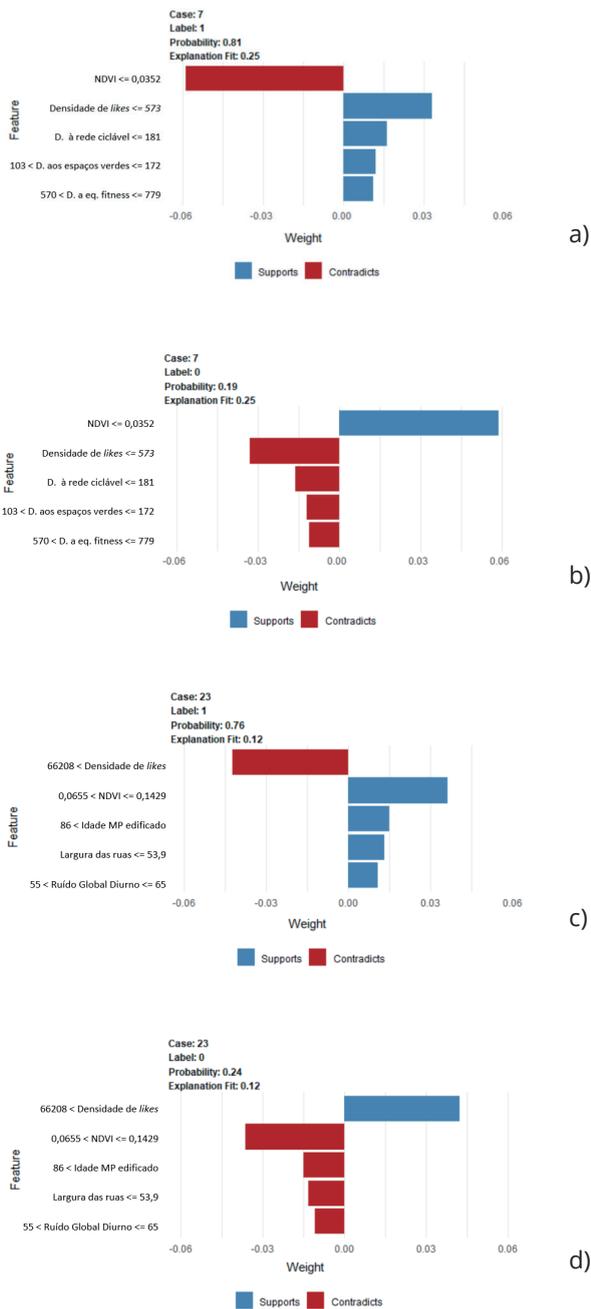


Figura 6. Interpretação do LIME para a instância 7 (a e b) e 23 (c e d) como sentimento positivo (a e c) e negativo (b e d).

5.4 Kernel SHAP

Para a interpretação local do modelo de caixa negra Xgboost utilizou-se o Kernel SHAP. O output do modelo indica-nos a contribuição das variáveis (i.e., ruído = 60) para predição do sentimento para cada subconjunto, através do cálculo da diferença entre a presença e ausência da variável para o subconjunto, efetuando a média ponderada de todas as contribuições (contribuição marginal) (Molnar, 2019). Representou-se graficamente duas observações individuais (7 e 23) dos dados do subconjunto local (Figura 7). Cada gráfico apresenta a probabilidade predita de cada observação como sentimento positivo e as variáveis que mais contribuíram para essa predição local. Indicando também se essas variáveis aumentam (suportam) ou diminuem (contradizem) a probabilidade de sentimento positivo.

A Figura 7-a) mostra a explicação Shapley da observação 7 com base nas 10 variáveis explicativas. Sendo que, a distância aos espaços verdes, energia solar, distância à rede ciclável, NDVI e ruído global são as variáveis que mais contribuíram para a predição do sentimento (positivo). As explicações mostram que valores de distância a espaços verdes de 122 metros suportam o sentimento positivo, contrariamente aos valores de NDVI de -0,0269 e de ruído global de 55 dB que contradizem.

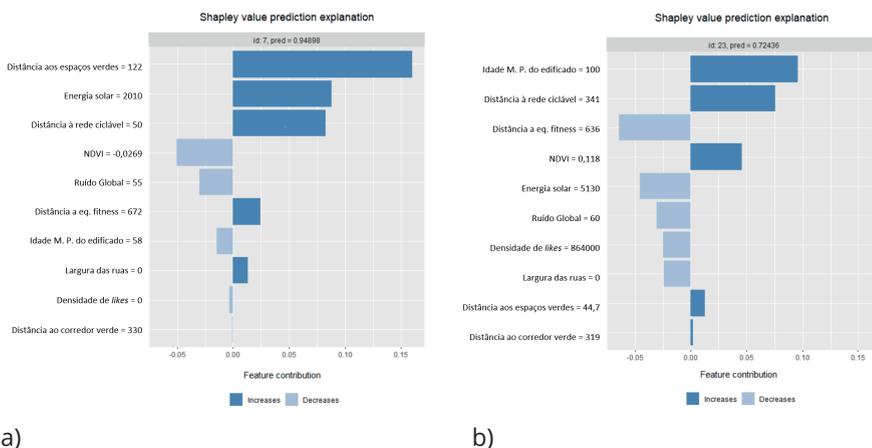


Figura 7. Interpretação Kernel SHAP para a instância 7 (a) e 23 (b) como sentimento positivo.

5.5 SHAP Summary Plot

O SHAP é um método de interpretação local, no entanto também é possível obter interpretações globais válidas agregando as predições individuais. Esta abordagem possibilita aprofundar um pouco mais a complexidade dos resultados do modelo preditivo, enquanto permite explorar as relações entre as features para a observação predita (Molnar, 2019).

A partir do modelo XGB, elaborou-se o Shap Summary Plot (Figura 8) para o conjunto de dados de sentimento. Pode-se verificar que valores mais elevados de distância aos equipamentos fitness estão associados a valores positivos de sentimento predito. Por outras palavras, as pessoas estão felizes se estiverem distantes dos equipamentos fitness. Quando o ruído é elevado, o SHAP value é positivo. Ou seja, as pessoas estão felizes quando o ruído é mais elevado, o que parece contraintuitivo. Contudo, isto poderá dever-se ao facto de a generalidade do ambiente urbano apresentar valores mais elevados de ruído.

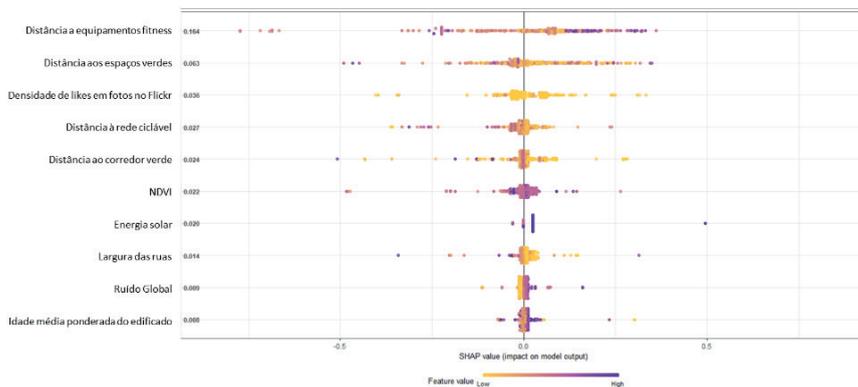


Figura 8. SHAP Summary plot.

6. Conclusões e limitações

Os dados do Twitter são um bom proxy para a análise de sentimento, uma vez que são atribuídos códigos aos utilizadores, assim sabemos quantos são, mas não sabemos quem são, não havendo problemas de confidencialidade. Os dados da localização são públicos nos casos em que o utilizador dá autorização, sendo possível em análises futuras seguir os utilizadores e ter uma perspetiva espacial e temporal do seu comportamento. A permissão para o acesso à localização do utilizador no momento da publicação para além de constituir uma vantagem no âmbito da proteção de dados, é uma desvantagem para a análise de sentimento. Isto, porque a maioria dos utilizadores não autoriza o acesso, acabando por não se considerar uma grande parte dos comentários para a análise.

A utilização de um léxico para a análise de sentimento é um processo facilmente replicável em R, pois basta invocar a biblioteca 'syuzhet' e escolher o método (NRC), não requerendo a execução de etapas intermedias, tais como a atribuição manual de rótulos para a criação de documentos de treino (e.g., abordagem de *machine learning* com método supervisionado).

As 10 variáveis mais influentes na explicação do sentimento, segundo o método global PFI, vão de encontro à Figura 1 de Hajrasoulih *et al.*, (2018), na qual são apresentadas as relações verificadas entre a forma urbana e o bem-estar

mental. Nomeadamente, a interação com o espaço verde, o design tradicional do bairro, idade e diversidade do edificado e o ruído, que se equiparam com as variáveis mais influentes, NDVI, distância ao corredor verde e espaços verdes, e idade média ponderada do edificado. Para além destas, entre as variáveis mais influentes (com uma pontuação de importância >50) na explicação do sentimento constam ainda a distância a equipamentos fitness, distância à rede ciclável, energia solar e a popularidade dos locais (densidade de likes em fotos no *Flickr*).

Os modelos agnósticos, principalmente o SHAP, desempenham um papel fundamental neste tipo de análise. Para além de explicar as predições individuais dos modelos de *machine learning*, também permite obter explicações globais válidas agregando as predições individuais. Respondendo à questão de partida, as variáveis que se relacionam mais com o sentimento, segundo este modelo, são a distância a equipamentos *fitness*, a distância aos espaços verdes, a popularidade dos locais e a distância à rede ciclável.

Este tipo de estudos são uma mais-valia para os planeadores, políticos, profissionais de saúde, que com a ajuda das interpretações do modelo de ML podem tomar decisões. Possibilitando desenvolver novas estratégias para reduzir as desigualdades e promover melhor a saúde e o bem-estar urbanos. Estes profissionais possuem um conhecimento prévio sobre o bem-estar em ambiente urbano, que podem utilizar para determinar se aceitam ou rejeitam uma predição.

7. Financiamento do projeto

The Portuguese Foundation for Science and Technology (FCT) supported this research via Grant No. 2022.11665.BD to Iúria Betco. The Centre for Geographical Studies - Universidade de Lisboa and FCT supported this research via Grant No. UIDB/00295/2020 + UIDP/00295/2020. We thank the GEOMODLAB—Laboratory for Remote Sensing, Geographical Analysis and Modelling—of the Centre for Geographical Studies (CEG) and Institute of Geography and Spatial Planning (IGOT) for providing the required equipment and software.

Bibliografia

- Abdelmoula, E., & Bouayed Abdelmoula, N. (2023). Promoting health and wellbeing through building the relationship with build and cultural heritage. *European Psychiatry*, 66(S1), S988–S988. DOI: <https://doi.org/10.1192/j.eurpsy.2023.2100>
- Abdul-Rahman, M., Chan, E. H. W., Wong, M. S., Irekponor, V. E., & Abdul-Rahman, M. O. (2021). A framework to simplify pre-processing location-based social media big data for sustainable urban planning and management. *Cities*, 109. DOI: <https://doi.org/10.1016/j.cities.2020.102986>
- Afriliana, N., Iswari, N. M. S., & Suryasari. (2022). Sentiment Analysis of User-Generated Content: A Bibliometric Analysis. *Journal of System and Management Sciences*, 12(6), 583–598. DOI: <https://doi.org/10.33168/JSMS.2022.0634>
- Aloufi, S., Zhu, S., & El Saddik, A. (2017). On the Prediction of Flickr Image Popularity by Analyzing Heterogeneous Social Sensory Data. *Sensors*, 17 (3). DOI: <https://doi.org/10.3390/s17030631>

- Ayala-Azcárraga, C., Diaz, D., & Zambrano, L. (2019). Characteristics of urban parks and their relation to user well-being. *Landscape and Urban Planning*, 189, 27–35. DOI: <https://doi.org/10.1016/j.landurbplan.2019.04.005>
- Benabbou, R., & Lee, H. (2019). Exploring the evolution of urban emotions in the City of Seoul using social media information. *International Journal of Knowledge-Based Development*, 10 (3), 232–248. DOI: <https://doi.org/10.1504/IJKBD.2019.103208>
- Cebeci, H. İ., Güner, S., Arslan, Y., & Aydemir, E. (2023). Barriers and drivers for biking: What can policymakers learn from social media analytics? *Journal of Transport and Health*, 28 (November 2022), 1–14. DOI: <https://doi.org/10.1016/j.jth.2022.101542>
- Chen, S., Liu, L., Chen, C., & Haase, D. (2022). The interaction between human demand and urban greenspace supply for promoting positive emotions with sentiment analysis from twitter. *Urban Forestry and Urban Greening*, 78. DOI: <https://doi.org/10.1016/j.ufug.2022.127763>
- Clark, C., Crumpler, C., & Notley, H. (2020). Evidence for environmental noise effects on health for the United Kingdom policy context: A systematic review of the effects of environmental noise on mental health, wellbeing, quality of life, cancer, dementia, birth, reproductive outcomes, and cognition. *International Journal of Environmental Research and Public Health*, 17 (2). DOI: <https://doi.org/10.3390/ijerph17020393>
- De la Riva, E. G., Ulrich, W., Batáry, P., Baudry, J., Beaumelle, L., Bucher, R., Čerevková, A., Felipe-Lucia, M. R., Gallé, R., Kesse-Guyot, E., Rembiałkowska, E., Rusch, A., Seufert, V., Stanley, D., & Birkhofer, K. (2023). From functional diversity to human well-being: A conceptual framework for agroecosystem sustainability. *Agricultural Systems*, 208 (November 2022). DOI: <https://doi.org/10.1016/j.agsy.2023.103659>
- Deci, E. L., & Ryan, R. M. (2008). *Hedonia, eudaimonia, and well-being: an introduction*. 1–11. DOI: <https://doi.org/10.1007/s10902-006-9018-1>
- Dohoo, I. R., Ducrot, C., Fourichon, C., Donald, A., & Hurnik, D. (1997). An overview of techniques for dealing with large numbers of independent variables in epidemiologic studies. *Preventive Veterinary Medicine*, 29(3), 221–239. DOI: [https://doi.org/10.1016/S0167-5877\(96\)01074-4](https://doi.org/10.1016/S0167-5877(96)01074-4)
- Du, G., Shin, K. J., & Managi, S. (2018). Variability in impact of air pollution on subjective well-being. *Atmospheric Environment*, 183 (April), 175–208. DOI: <https://doi.org/10.1016/j.atmosenv.2018.04.018>
- Eddolls, W. T. B., McNarry, M. A., Lester, L., Winn, C. O. N., Stratton, G., & Mackintosh, K. A. (2018). The association between physical activity, fitness and body mass index on mental well-being and quality of life in adolescents. *Quality of Life Research*, 27 (9), 2313–2320. DOI: <https://doi.org/10.1007/s11136-018-1915-3>
- Elshawi, R., Al-Mallah, M. H., & Sakr, S. (2019). On the interpretability of machine learning-based model for predicting hypertension. *BMC Medical Informatics and Decision Making*, 19 (1), 146. DOI: <https://doi.org/10.1186/s12911-019-0874-0>
- Fathi, S., Sajadzadeh, H., Sheshkal, F. M., Aram, F., Pinter, G., Felde, I., & Mosavi, A. (2020). The role of urban morphology design on enhancing physical activity and public health. *International Journal of Environmental Research and Public Health*, 17 (7), 1–29. DOI: <https://doi.org/10.3390/ijerph17072359>
- Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29 (5), 1189–1232.

- Gai, Z., Fan, C., Shen, S., Ge, Y., Shi, Z., Li, S., Zhang, Y., & Cao, Y. (2022). Using Social Media Data to Explore Urban Land Value and Sentiment Inequality: A Case Study of Xiamen, China. *Wireless Communications and Mobile Computing*, 2022. DOI: <https://doi.org/10.1155/2022/1456382>
- Galinha, I., & Ribeiro, J. L. P. (2005). História e Evolução do Conceito de Bem-Estar Subjectivo, *Psicologia, Saúde & Doenças*, 6 (2), 203-214.
- Gao, Y., Chen, Y., Mu, L., Gong, S., Zhang, P., & Liu, Y. (2022). Measuring urban sentiments from social media data: a dual-polarity metric approach. *Journal of Geographical Systems*, 24 (2), 199–221. DOI: <https://doi.org/10.1007/s10109-021-00369-z>
- Ghahramani, M., Galle, N. J., Duarte, F., Ratti, C., & Pilla, F. (2021). Leveraging artificial intelligence to analyze citizens' opinions on urban green space. *City and Environment Interactions*, 10 (February), 100058. DOI: <https://doi.org/10.1016/j.cacint.2021.100058>
- Gjerde, M., & Vale, B. (2022). An examination of people's preferences for buildings and streetscapes in New Zealand. *Australian Planner*, 58 (1-2), 36-48. DOI: <https://doi.org/10.1080/07293682.2022.2029922>
- Hajrasoulih, A., del Rio, V., Francis, J., & Edmondson, J. (2018). Urban form and mental wellbeing. *Journal of Urban Design and Mental Health*, 5 (1), 10. DOI: <https://www.urbandesignmentalhealth.com/journal-5---urban-form-and-mental-wellbeing.html>
- Harrison, P. L., Pope, J. E., Coberley, C. R., & Rula, E. Y. (2012). Evaluation of the relationship between individual well-being and future health care utilization and cost. *Population Health Management*, 15(6), 325–330. DOI: <https://doi.org/10.1089/pop.2011.0089>
- He, H., Sun, R., Li, J., & Li, W. (2023). Urban landscape and climate affect residents' sentiments based on big data. *Applied Geography*, 152. DOI: <https://doi.org/10.1016/j.apgeog.2023.102902>
- Hidaka, B. H. (2012). Depression as a disease of modernity: explanations for increasing prevalence. *National Institutes of Health*, 140 (3), 205–214. DOI: <https://doi.org/10.1016/j.jad.2011.12.036>. Depression
- Huang, S., Zhu, J., Zhai, K., Wang, Y., Wei, H., Xu, Z., & Gu, X. (2022). Do Emotional Perceptions of Visible Greeneries Rely on the Largeness of Green Space? A Verification in Nanchang, China. *Forests*, 13 (8). DOI: <https://doi.org/10.3390/f13081192>
- Huppert, F. A. (2009). *Psychological Well-being: Evidence Regarding its Causes and Consequences*. 1 (2), 137-164. DOI: <https://doi.org/10.1111/j.1758-0854.2009.01008.x>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *Springer Texts in Statistics An Introduction to Statistical Learning* (Springer (ed.); 7th ed.). DOI: <https://doi.org/10.1007/978-1-4614-7138-7>
- Jiao, Y., Li, C., & Lin, Y. (2021). Can urban environmental problems be accurately identified? A complaint text mining method. *Applied Sciences (Switzerland)*, 11(9). DOI: <https://doi.org/10.3390/app11094087>
- Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality & Quantity*, 52 (4), 1957-1976. DOI: <https://doi.org/10.1007/s11135-017-0584-6>

- Keyes, C. L. E. M. (1998). Social well-being. *Social Psychology Quarterly*, 61 (2), 121–140.
- Lai, S., & Deal, B. (2023). Parks, Green Space, and Happiness: A Spatially Specific Sentiment Analysis Using Microblogs in Shanghai, China. *Sustainability (Switzerland)*, 15 (1). DOI: <https://doi.org/10.3390/su15010146>
- Lawton, R. N., & Fujiwara, D. (2016). Living with aircraft noise: Airport proximity, aviation noise and subjective wellbeing in England. *Transportation Research Part D: Transport and Environment*, 42, 104-118.
DOI: <https://doi.org/10.1016/j.trd.2015.11.002>
- Leyland, L. A., Spencer, B., Beale, N., Jones, T., & van Reekum, C. M. (2019). The effect of cycling on cognitive function and well-being in older adults. *PLoS ONE*, 14 (2), 1-17. DOI: <https://doi.org/10.1371/journal.pone.0211779>
- Lin, F.-J. (2008). Solving Multicollinearity in the Process of Fitting Regression Model Using the Nested Estimate Procedure. *Quality & Quantity*, 42 (3), 417–426.
DOI: <https://doi.org/10.1007/s11135-006-9055-1>
- Lingjærde, C., Jullum, M., & Sellereite, N. (2021). *Explaining individual machine learning predictions with Shapley values*.
- Lipovčan, L. K., Brajša-Žganec, A., & Poljanec-Borić, A. S. (2014). What is good for tourists should be good for residents too: The relationship between the quality of the touristic offer and subjective well-being of residents. *Tourism Analysis*, 19 (6), 719–730. DOI: <https://doi.org/10.3727/108354214X14146846679448>
- Lorenc, T., Clayton, S., Neary, D., Whitehead, M., Petticrew, M., Thomson, H., Cummins, S., Sowden, A., & Renton, A. (2012). Crime, fear of crime, environment, and mental health and wellbeing: Mapping review of theories and causal pathways. *Health and Place*, 18 (4), 757-765.
DOI: <https://doi.org/10.1016/j.healthplace.2012.04.001>
- Lundberg, S. M., & Lee, S. (2017). *A Unified Approach to Interpreting Model Predictions. Section 2*, 1-10.
- Marsillo, L., Suntorachai, N., Karthikeyan, K. N., Voinova, N., Khairallah, L., & Chronis, A. (2022). Context Decoder Measuring urban quality through artificial intelligence. *Proceedings of the International Conference on Education and Research in Computer Aided Architectural Design in Europe*, 2 (September), 237-246.
DOI: <https://doi.org/10.52842/conf.eacaade.2022.2.237>
- Martín-Martín, J. M., Prados-Castillo, J. F., Jiménez Aguilera, J. de D., & Porrás González, E. (2023). Interferences generated on the well-being of local communities by the activity of online platforms for tourist accommodation. *Journal of Sustainable Tourism*, 31(2), 483–503. <https://doi.org/10.1080/09669582.2020.1861455>
- Mayor, J., & Coleman, B. (2011). *The Social and Emotional Benefits of Good Street Design*. August.
- McGillivray, M., & Clarke, M. (2006). *Understanding human well-being*. United Nations University.
- Mohammad, S., & Turney, P. (2015). *NRC Word-Emotion Association Lexicon (aka EmoLex)*. <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- Molnar, C. (2019). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Independently published.
- Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography*, 37(2), 191–203.
DOI: <https://doi.org/https://doi.org/10.1111/j.1600-0587.2013.00205.x>

- Reeves, W. C., Strine, T. W., Pratt, L. A., Thompson, W., Ahluwalia, I., Dhingra, S. S., McKnight-Eily, L. R., Harrison, L., D'Angelo, D. V., Williams, L., Morrow, B., Gould, D., & Safran, M. A. (2011). *Mental Illness Surveillance Among Adults in the United States*. 60.
- Reilly, S., Nolan, C., & Monckton, L. (2018). *Wellbeing and the Historic Environment: Threats, Issues and Opportunities for the Historic Environment*. 01.
- Ristea, A., Leitner, M., Resch, B., & Stratmann, J. (2021). Applying spatial video geonarratives and physiological measurements to explore perceived safety in baton rouge, louisiana. *International Journal of Environmental Research and Public Health*, 18 (3), 1-20. DOI: <https://doi.org/10.3390/ijerph18031284>
- Romice, O., Thwaites, K., Porta, S., Greaves, M., Barbour, G., & Pasino, P. (2016). City form and wellbeing. In O. Fleury-Bahi, Ghoulane, Pol, Enric, Navarro (Ed.), *The Handbook of Environmental Psychology and Quality of Life Research* (Issue January, pp. 241–273). Springer. https://doi.org/10.1007/978-3-319-31416-7_14
- Salami, D., Sousa, Carla A., Martins, M. do R. O., & Capinha, C. (2020). Predicting dengue importation into Europe, using machine learning and model-agnostic methods. *Scientific Reports*, 1-14. DOI: <https://doi.org/10.1038/s41598-020-66650-1>
- Santana, P., Nogueira, H., & Santos, R. (2007). *Melhorar a saúde na Amadora Intervindo no Ambiente Físico e Social* (pp. 239–254).
- Schwartz, A. J., Dodds, P. S., O'Neil-Dunne, J. P. M., Danforth, C. M., & Ricketts, T. H. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. *People and Nature*, 1 (4), 476–485. DOI: <https://doi.org/10.1002/pan3.10045>
- Shayaa, S., Jaafar, N. I., Bahri, S., Sulaiman, A., Wai, P. S., Chung, Y. W., Piprani, A. Z., & Al-garadi, M. A. (2018). Sentiment Analysis of Big Data : Methods , Applications , and Open Challenges. *IEEE Access*, 6 (December 2019), 37807–37827. DOI: <https://doi.org/10.1109/ACCESS.2018.2851311>
- Stellacci, S., & Moro, S. (2022). Travellers' perspectives on historic squares and railway stations in Italian heritage cities revealed through sentiment analysis. *Journal of Urban Design*. DOI: <https://doi.org/10.1080/13574809.2022.2097903>
- Tao, J., Yang, M., & Wu, J. (2022). Coupling Coordination Evaluation of Lakefront Landscape Spatial Quality and Public Sentiment. *Land*, 11(6). <https://doi.org/10.3390/land11060865>
- United Nations. (2015). *World Urbanization Prospects: The 2014 Revision*. ST/ESA/SER.A/366
- Viana, C. M., Santos, M., Freire, D., Abrantes, P., & Rocha, J. (2021). *Evaluation of the factors explaining the use of agricultural land: A machine learning and model-agnostic approach*. 131. DOI: <https://doi.org/10.1016/j.ecolind.2021.108200>
- Völker, S., & Kistemann, T. (2011). The impact of blue space on human health and well-being - Salutogenetic health effects of inland surface waters: A review. *International Journal of Hygiene and Environmental Health*, 214 (6), 449-460. DOI: <https://doi.org/10.1016/j.ijheh.2011.05.001>
- Wang, J., Wu, Y., & Xia, Y. (2022). A study on the distribution of 5A attractions and the Spatio-temporal characteristics of tourists' emotions in China. *International Conference on Geoinformatics*, 2022-Augus, 1-7. DOI: <https://doi.org/10.1109/Geoinformatics57846.2022.9963808>
- Wang, R., Yang, B., Yao, Y., Bloom, M. S., Feng, Z., Yuan, Y., Zhang, J., Liu, P., Wu, W., Lu, Y., Baranyi, G., Wu, R., Liu, Y., & Dong, G. (2020). Residential greenness, air

pollution and psychological well-being among urban residents in Guangzhou, China. *Science of the Total Environment*, 711, 1-12.

DOI: <https://doi.org/10.1016/j.scitotenv.2019.134843>

White, M. P., Elliott, L. R., Gascon, M., Roberts, B., & Fleming, L. E. (2020). Blue space, health and well-being: A narrative overview and synthesis of potential benefits. *Environmental Research*, 191 (August), 110169.

DOI: <https://doi.org/10.1016/j.envres.2020.110169>

Xiong, Y., He, Y., Huang, H., Yu, C., & Jing, X. (2020). Air Quality Statistics and Prediction Based on Urban Agglomerations and Sentiment Analysis of People Under Different Pollutants. *Lecture Notes in Electrical Engineering*, 628 LNEE, 78-87. DOI: https://doi.org/10.1007/978-981-15-4163-6_10

Ye, Z., Li, R., & Wu, J. (2022). Dynamic Demand Evaluation of COVID-19 Medical Facilities in Wuhan Based on Public Sentiment. *International Journal of Environmental Research and Public Health*, 19(12). <https://doi.org/10.3390/ijerph19127045>

Yuen, H. K., & Jenkins, G. R. (2020). Factors associated with changes in subjective well-being immediately after urban park visit. *International Journal of Environmental Health Research*, 30(2), 134–145.

DOI: <https://doi.org/10.1080/09603123.2019.1577368>

Zhang, J., Li, X., Bao, T., Li, Z., Liu, C., & Xu, Y. (2021). Linking demographic factors, land use, ecosystem services, and human well-being: Insights from an sandy landscape, uxin in Inner Mongolia, China. *Sustainability (Switzerland)*, 13 (9).

DOI: <https://doi.org/10.3390/su13094847>