

A IMPORTÂNCIA DO ESPAÇO URBANO PARA O SENTIMENTO: ANÁLISE BASEADA EM REDES SOCIAIS

Betco, Iuria¹; David, Melissa¹; Rocha, Jorge^{1,2} Nogueira, Paulo^{2,3}

¹Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa,
betcoiuria@gmail.com

¹Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa,
melissa.m.d.s.17@gmail.com

²Laboratório Associado Terra, jorge.rocha@campus.ul.pt

³ Faculdade de Medicina de Lisboa, Laboratório de Biomatemática - Instituto de Medicina
Preventiva e Saúde pública, pnogueira@medicina.ulisboa.pt

RESUMO: O ambiente urbano há muito que é reconhecido como um importante determinante da saúde e do bem-estar. Historicamente, este centrou-se na poluição, na doença e na sobrepopulação, mas existem agora provas crescentes dos impactes do ambiente urbano físico na saúde e bem-estar, tanto positivos como negativos. Este estudo tem o intuito de compreender o impacte que o ambiente urbano tem na felicidade dos indivíduos. Para tal, efetuou-se uma análise do bem-estar com recurso à rede social Twitter que permitiu determinar quais os aspetos da experiência no ambiente urbano que afetam o bem-estar dos portugueses na cidade de Lisboa no ano de 2019. De modo a determinar quais as variáveis que influenciam o sentimento positivamente e negativamente aplicou-se *machine learning* (ML). Este subcampo da inteligência artificial (IA) é reconhecido por alcançar resultados de exatidão superiores aos métodos tradicionais, apresentando ainda vantagens, como a capacidade de lidar com dados de diferentes tipos, estruturas e quantidades, não sendo sensíveis à escala das variáveis, possibilitando combinar dados de diversas fontes para modelar relações não lineares complexas que descrevam a variação do sentimento em ambientes urbanos. Com este estudo foi possível concluir que os portugueses são felizes em espaços associados ao lazer e ao consumo, tais como, museus, locais de eventos, jardins, centros comerciais, lojas e restaurantes.

PALAVRAS CHAVE: Bem-estar; redes sociais; análise de sentimento; morfologia urbana; *machine learning*.

ABSTRACT: The urban environment has long been recognized as an important determinant of health and well-being. Historically this has focused on pollution, disease, and overcrowding, but there is now growing evidence of the impacts of the physical urban environment on health and well-being, both positive and negative. This study aims to understand the impact that the urban environment has on individuals' happiness. To this end, a well-being analysis was carried out using the social network Twitter, which allowed us to determine which aspects of the urban environment experience affect the well-being of Portuguese people in the city of Lisbon in 2019. To determine which variables influence the feeling positively and negatively, machine learning (ML) was applied. This subfield of artificial intelligence (AI) is recognized for achieving higher accuracy results than traditional methods, also presenting advantages, such as the ability to deal with data of different types, structures, and quantities, not being sensitive to the scale of the variables, making it possible to combine data from various sources to model complex nonlinear relationships that describe the variation of sentiment in urban environments. With this study, it was possible to conclude that the Portuguese are happy in spaces associated with leisure and consumption, such as museums, event venues, gardens, shopping centers, stores, and restaurants.

KEYWORDS: Wellbeing; social networks; sentiment analysis; urban morphology; machine learning.

1. INTRODUÇÃO

Os planeadores e profissionais de saúde têm demonstrado preocupação com o impacto dos ambientes urbanos no bem-estar mental dos residentes (Hidaka, 2012). Embora a qualidade de vida em geral tenha melhorado em todo o mundo, os problemas de saúde mental aumentaram, o que pode estar associado ao estilo de vida urbano (Reeves et al., 2011). De facto, algumas das doenças mentais mais prevalentes parecem ser causadas pelos elevados níveis de stress despoletados pela vida nas cidades. Estudos recentes constataram que, em moradores urbanos, o risco de desenvolver transtornos mentais é 38% maior (39% maior para transtornos do

humor, 21% maior para transtornos de ansiedade) do que em pessoas que vivem em áreas rurais (Peen et al., 2010).

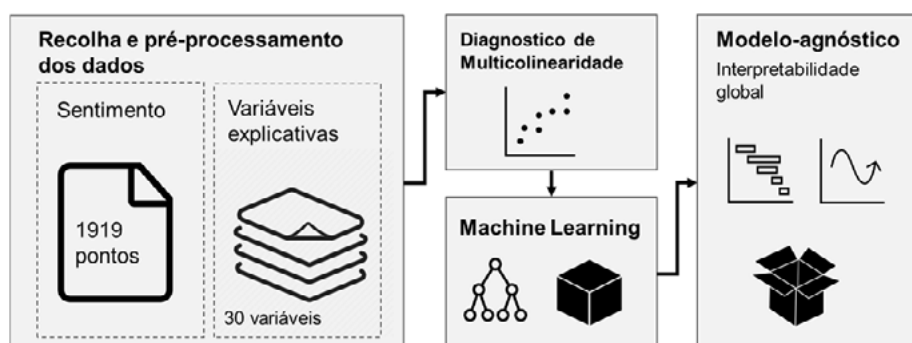
Apesar de ter sido realizada uma grande quantidade de investigações sobre as potenciais associações entre as áreas urbanas e a saúde ou bem-estar, há falta de abordagens interdisciplinares que destaquem a complexidade das estruturas e dinâmicas urbanas e a sua possível influência na saúde e bem-estar urbanos (Krefis et al., 2018). Neste sentido, é cada vez mais importante compreender o impacto da cidade na saúde humana (UN Habitat, 2013), sendo 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, permite o conhecimento acerca do quanto o bem-estar individual é influenciado pelo ambiente construído envolvente (Romice et al., 2016), através do conteúdo gerado pelo utilizador nas redes sociais. O conteúdo gerado pelos internautas nas redes sociais fornece informações sobre os utilizadores e o seu ambiente envolvente, uma vez que permite aceder às suas preferências, opiniões e interações (Aloufi et al., 2017).

2. METODOLOGIA

Este estudo explora a utilização de um modelo de *Machine Learning* treinado com o algoritmo *Ranfom Forest* juntamente com uma abordagem de modelo-agnóstico, de modo a aumentar a compreensão das condições dos ambientes urbanos que podem explicar o bem-estar subjetivo (Figura 1).

Figura 1. Esquema metodológico.



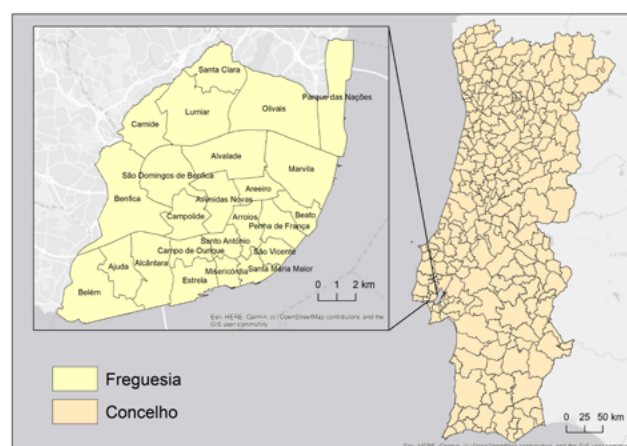
Atualmente, existem exemplos de modelos agnósticos interpretáveis globais e locais. Dos globais utilizou-se o *Permutation Feature Importance* (PFI) e o *Partial Dependence Plot* (PDPs) (Molnar, 2019). O método PFI permitiu compreender quais as variáveis que contribuíram para os resultados subjacentes ao modelo ML e quantificar a sua importância. Este método mede o aumento do erro de predição do modelo após os valores da variável serem permutados (Molnar, 2019; Winkler et al., 2015). O método PDP representa a relação global das variáveis explicativas com a variável dependente sob a forma de gráfico, impondo que todas as ocorrências variem dentro da mesma gama de valores e meçam o efeito marginal ou médio para esse mesmo valor na resposta do modelo (Molnar, 2019).

De modo a identificar automaticamente os sentimentos (bem-estar) expressos nos comentários publicados pelos utilizadores da rede social *Twitter* utilizou-se o léxico de sentimentos *EmoLex*, que consiste numa lista de palavras manualmente rotuladas com polaridade positiva (e.g., divertido) e negativa (e.g., triste).

3. ÁREA DE ESTUDO

A cidade de Lisboa, capital de Portugal, situa-se junto ao estuário do rio Tejo e apresenta uma área de 86,83 km². O município de Lisboa subdivide-se em 24 freguesias representadas na figura (Figura 2).

Figura 2. Freguesias do município de Lisboa.



Em termos populacionais, o município de Lisboa alberga 509 515 habitantes (INE, 2021), apresentando uma densidade populacional de 5 868 habitantes por km². No que diz respeito às

estatísticas da procura turística, o município de Lisboa é o destino com maior procura a nível nacional (INE, 2019b). Apresenta uma ampla gama de atividades turísticas, destacando-se pela 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. RESULTADOS E DISCUSSÃO

4.1. Análise de sentimento

Os mapas gerados a partir dos resultados obtidos da análise de sentimento permitiram averiguar quais as áreas da cidade de Lisboa com maior pontuação de palavras com associação ao sentimento positivo (Figura 3) e negativo (Figura 4).

Figura 3. Pontuação de palavras do Twitter com associação ao sentimento positivo no município de Lisboa.

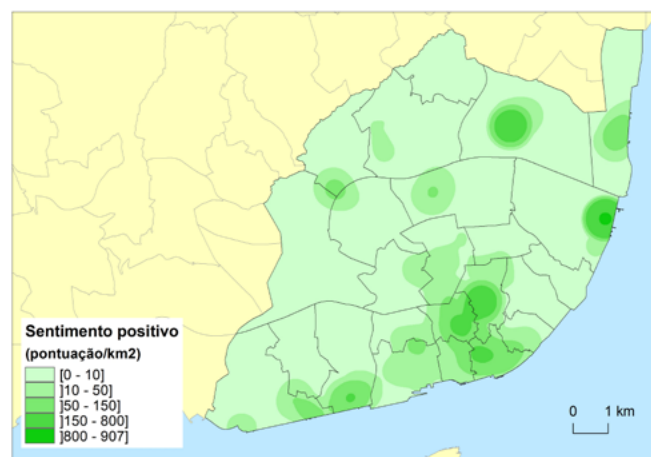
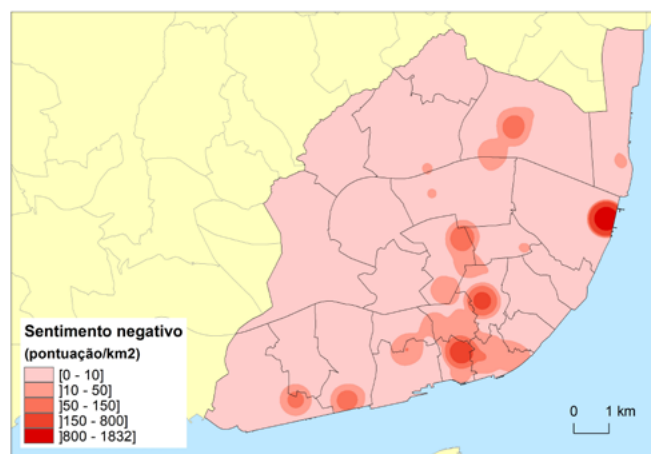


Figura 4. Pontuação de palavras do Twitter com associação ao sentimento negativo no município de Lisboa.



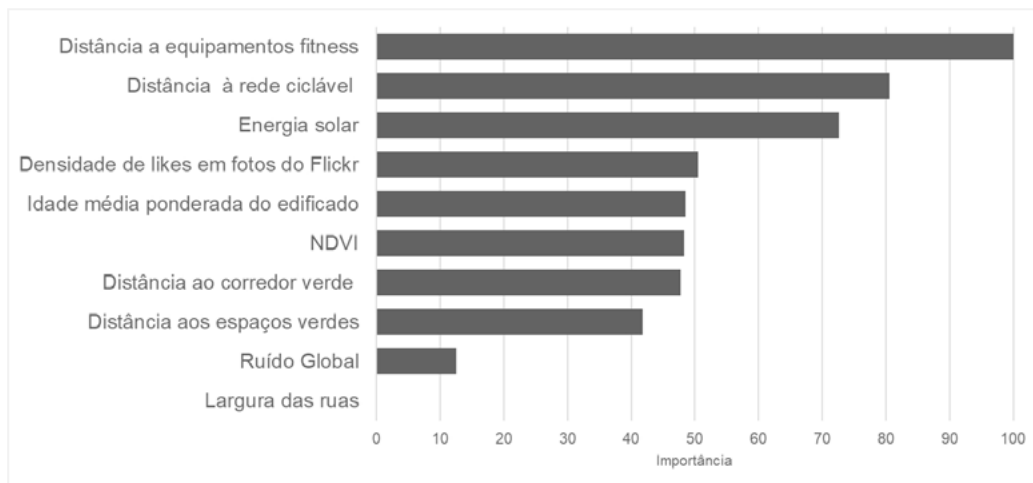
Através destas figuras (Figura 3 e Figura 4) constatou-se que as palavras com associação a sentimentos negativos apresentam uma maior concentração na cidade de Lisboa, contrariamente às palavras com associação a sentimentos positivos que se encontram mais dispersas.

A partir da Figura 3 é possível verificar que as áreas com elevada pontuação de palavras com associação a sentimento positivo ([150 - 800]) abrange as lojas e restaurantes da Avenida da Liberdade, a Praça Dom Pedro IV (Praça do Rossio), o Elevador de Santa Justa, os Armazéns do Chiado, a Rua Augusta e o Aeroporto Humberto Delgado. Com pontuação moderada ([150 - 800]) encontram-se abrangidos o MAAT, a Doca de Santo Amaro, o Centro de Congressos de Lisboa, o Village Underground Lisboa, o LxFactory, o SUD Lisboa, o Jardim Guerra Junqueiro (Jardim da Estrela), o Centro Comercial do Colombo, o Centro Comercial Vasco da Gama, o Altice Arena e a Feira Internacional de Lisboa (FIL). A generalidade dos locais (eg., Bairro Alto, Campo pequeno, Torre de Belém e Cristina Guerra) identificados como tendo uma elevada pontuação de palavras com associação ao sentimento negativo (Figura 4), deve-se ao facto de o léxico *EmoLex* não ser capaz de identificar o contexto em que a palavra aparece, acabando por lhe atribuir uma pontuação negativa (eg., alto, pequeno, torre e guerra).

4.1. Abordagem de modelo-agnóstico

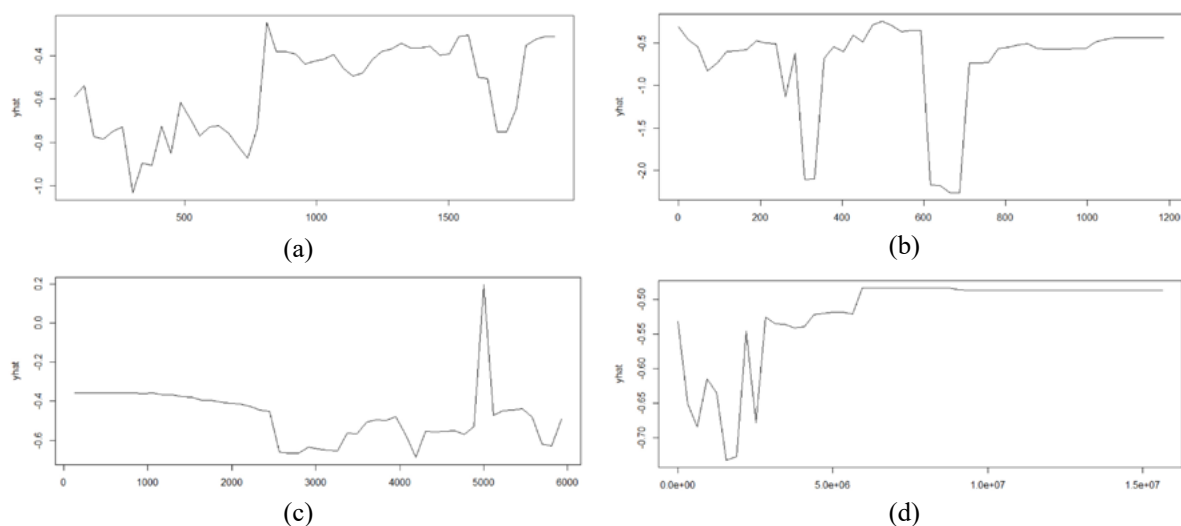
Os *outputs* da modelação *Random Forest* foram examinados utilizando o método PFI para a interpretação global. A Figura 5, apresenta as 10 variáveis mais influentes na explicação do sentimento.

Figura 5. As 10 variáveis mais influentes na explicação do sentimento segundo o método global PFI.



As variáveis mais influentes (com uma pontuação de importância > 50) na explicação do sentimento foram a distância a equipamentos *fitness*, distância à rede ciclável, energia solar e a popularidade dos locais (densidade de *likes* em fotos do *Flickr*). O método PDP foi também utilizado para a interpretação global do modelo, a Figura 6 apresenta as curvas de resposta para as primeiras quatro variáveis mais influentes e o quanto se traduz no incremento positivo ou negativo do sentimento.

Figura 6. Curvas de resposta das primeiras 4 variáveis mais influentes no incremento do sentimento segundo o método global PDP.



Analisando cada gráfico em particular, pode-se verificar que: (Figura 6-a) a proximidade aos equipamentos *fitness* não influencia positivamente o sentimento, verificando-se uma preferência por distâncias maiores a estes; (Figura 6-b) a proximidade à rede ciclável influencia de um modo geral positivamente o sentimento; (Figura 6-c) valores muito baixos

de energia solar traduzem-se em valores relativamente baixos de sentimento, o pico nos 5000 kW diários que se traduz numa preferência por espaços com maior exposição solar; (Figura 6-d) a popularidade dos locais (*likes/km²*) influencia positivamente o sentimento, verificando-se também uma preferência por locais mais populares.

5. CONCLUSÃO

Pode-se concluir que os portugueses são felizes em espaços associados ao lazer e ao consumo, tais como, museus, locais de eventos, jardins, centros comerciais, lojas e restaurantes. A generalidade dos locais identificados como tendo uma elevada pontuação de palavras com associação ao sentimento negativo deve-se ao facto de o léxico *EmoLex* não ser capaz de identificar o contexto em que a palavra aparece, acabando por lhe atribuir uma pontuação negativa, podendo levar a um enviesamento dos resultados. Portanto, a escolha deste léxico para a análise do sentimento, poderá não ter sido a opção mais viável para o resultado pretendido, devido ao formato e natureza dos comentários do Twitter. As variáveis mais influentes na explicação do sentimento na cidade de Lisboa são a distância aos equipamentos *fitness*, distância à rede ciclável, energia solar, popularidade dos locais. Verifica-se uma preferência pela proximidade à rede ciclável, contrariamente aos equipamentos *fitness* aos quais se verifica preferência por distâncias superiores. Também a elevada exposição à luz solar e a popularidade dos locais se afirmam como preferência dos utilizadores da cidade de Lisboa.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Aloufi, S., Zhu, S., & El Saddik, A. (2017). On the Prediction of Flickr Image Popularity by Analyzing Heterogeneous Social Sensory Data. In *Sensors* (Vol. 17, Issue 3).
<https://doi.org/10.3390/s17030631>
- 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.
<https://doi.org/10.3390/ijerph17072359>
- Hidaka, B. H. (2012). Depression as a disease of modernity: explanations for increasing prevalence. *National Institutes of Health*, 140(3), 205–214.
<https://doi.org/10.1016/j.jad.2011.12.036>. Depression
- Instituto Nacional de Estatística (INE). (2019). *Hóspedes (N.º) nos estabelecimentos de alojamento turístico por Localização geográfica (NUTS - 2013) e Tipo (alojamento turístico); Anual*.
- Instituto Nacional de Estatística (INE). (2021). *População residente (N.º) por Local de residência (resultados preliminares Censos2021) e Sexo; Decenal*.

- Krefis, A., Augustin, M., Schlünzen, K., Obenbrügge, J., & Augustin, J. (2018). How Does the Urban Environment Affect Health and Well-Being? A Systematic Review. *Urban Science*, 2(1), 21. <https://doi.org/10.3390/urbansci2010021>
- Molnar, C. (2019). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*.
- Observatório Turismo de Lisboa. (2019). *Inquérito às Atividades dos Turistas e Informação: Região de Lisboa*.
- Peen, J., Schoevers, R. A., Beekman, A. T., & Dekker, J. (2010). The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatrica Scandinavica*, 121(2), 84–93. <https://doi.org/https://doi.org/10.1111/j.1600-0447.2009.01438.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.
- Romice, O., Thwaites, K., Porta, S., Greaves, M., Barbour, G., & Pasino, P. (2016). City form and wellbeing. In O. Fleury-Bahi, Ghozlane, 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
- UN Habitat. (2013). *State of the World's Cities 2012: Prosperity of Cities*.
- Winkler, A. M., Webster, M. A., Vidaurre, D., Nichols, T. E., & Smith, S. M. (2015). Multi-level block permutation. *NeuroImage*, 123, 253–268. <https://doi.org/https://doi.org/10.1016/j.neuroimage.2015.05.092>