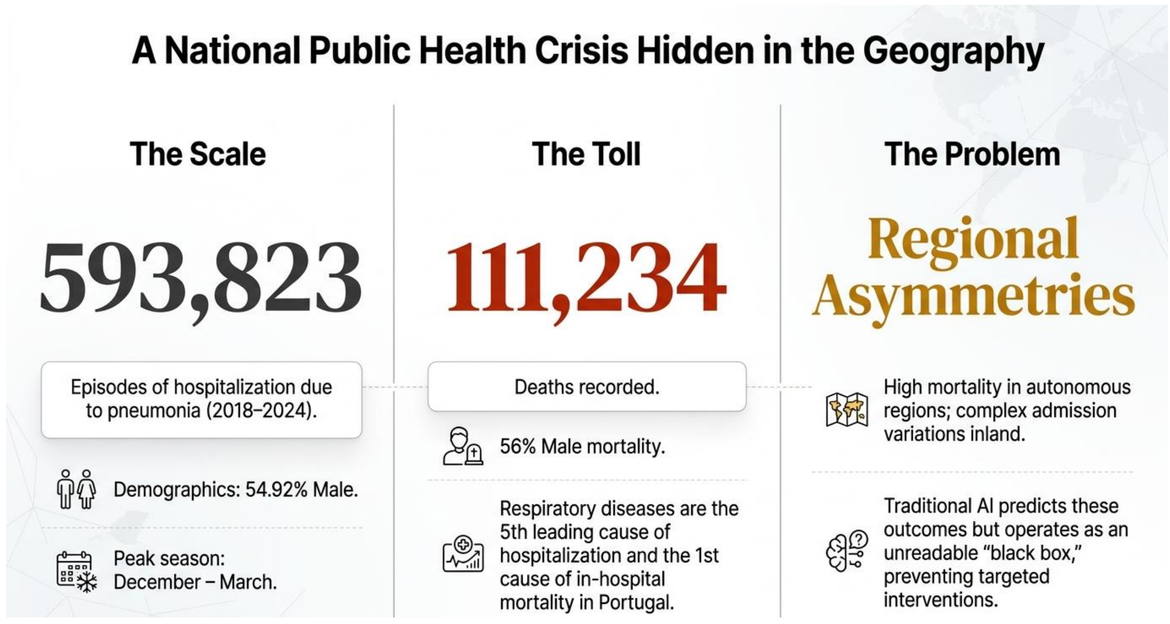
Understanding the Spatial Determinants and Geographical Inequalities of Community-Acquired Pneumonia in Portugal through AI and xAI models

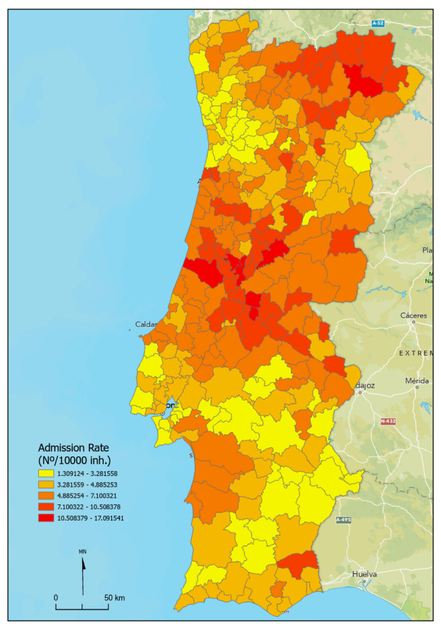
Jorge Rocha*, Iuria Betco, Eduardo Gomes and Cláudia M. Viana
* Centre of Geographical Studies, Institute of Geography and Spatial Planning, University of Lisbon; jorge.rocha@edu.ulisboa.pt

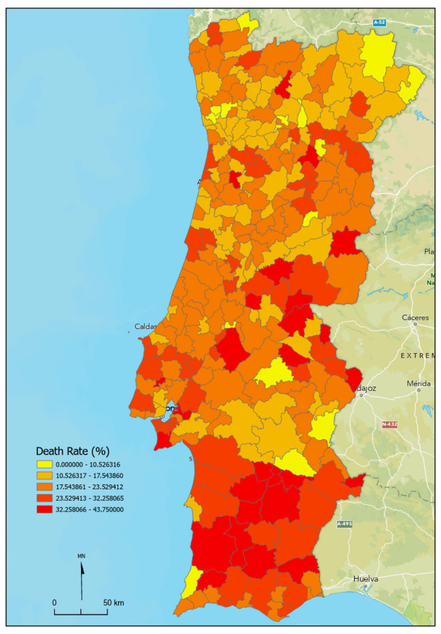
GeoAI 2026

INTRODUCTION

Community-acquired pneumonia (CAP) is a common and potentially serious infectious disease with multiple causative agents. Understanding the risk factors associated with different pathogens is fundamental to defining more effective treatment and clinical management strategies. The incidence of CAP increases with age, with higher rates observed among people over 65 years old. Environmental factors play a significant role in the development of CAP. Exposure to tobacco smoke, air pollution, and overcrowded housing conditions increases the risk of pneumonia. These determinants often interact with individual variables (such as comorbidities or vaccination status), reinforcing the need for integrated public health approaches.

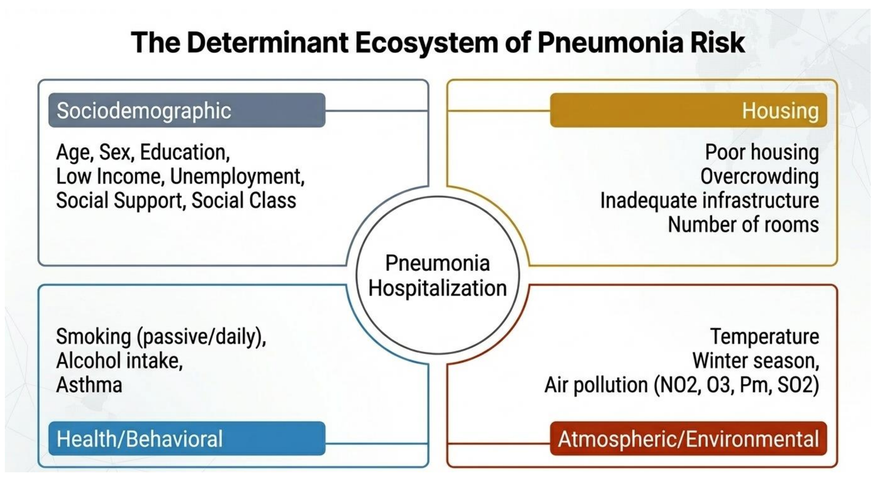




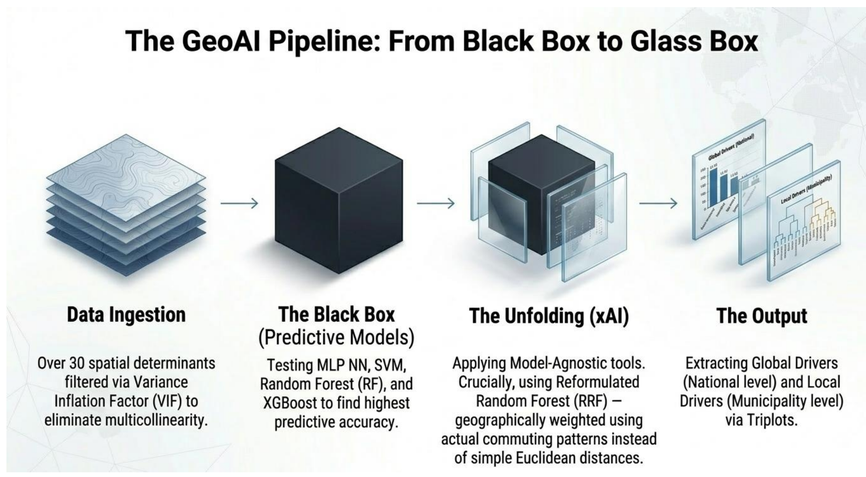


PROBLEM

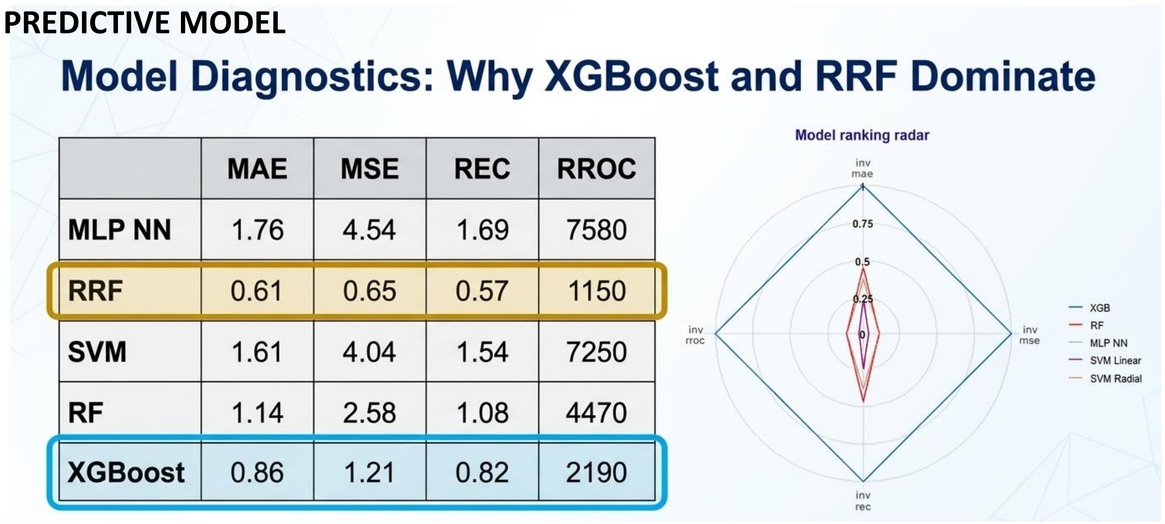
Geospatial artificial intelligence (GeoAI) models have demonstrated utility in predicting adverse outcomes in patients with community-acquired pneumonia (CAP), such as the need for hospitalization and mortality. In terms of predictive accuracy, these models have outperformed traditional clinical risk stratification tools, such as the Pneumonia Severity Index (PSI). However, the application of GeoAI models, particularly machine learning (ML) models, to the diagnosis and prognosis of CAP faces known limitations: it requires reliable and balanced databases, is subject to potential training data biases, and often involves "black box" models, whose interpretability is reduced for clinicians and decision-makers (Viana & Rocha, 2025).



In our work, we analyzed a set of over 30 spatial determinants at the municipal level in mainland Portugal (selected based on a literature review). After eliminating multicollinearity problems using the Variance Inflation Factor (VIF), we tested different approaches. In the end, RRF and XGBoost showed the best predictive performance.

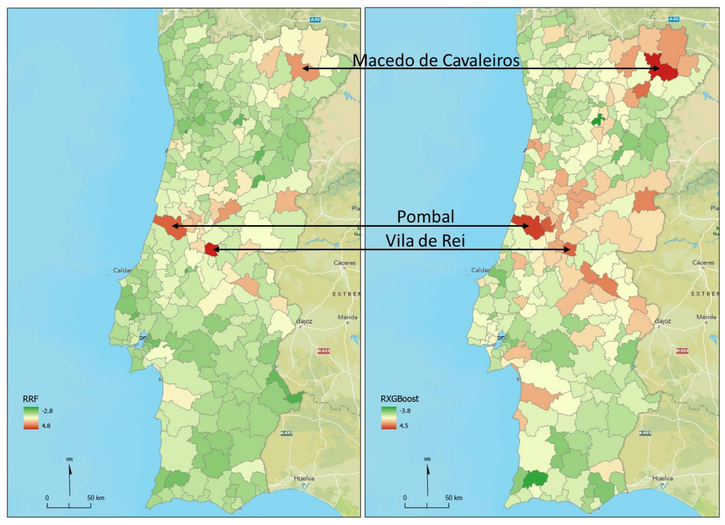


PREDICTIVE MODEL



	MAE	MSE	REC	RROC
MLP NN	1.76	4.54	1.69	7580
RRF	0.61	0.65	0.57	1150
SVM	1.61	4.04	1.54	7250
RF	1.14	2.58	1.08	4470
XGBoost	0.86	1.21	0.82	2190

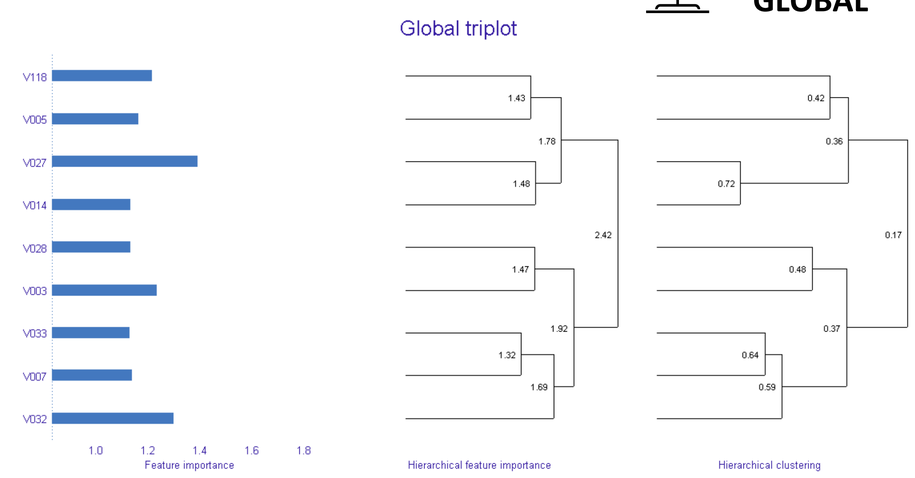
Model Residuals

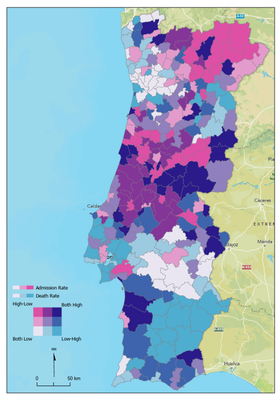


To interpret the results, we used explainable AI (xAI), that is, model-agnostic methods, evaluating the importance of each variable (Betco & Rocha, 2024; Viana et al., 2021). In particular, we used Triplot, which considers the hierarchical correlation structure between variables, allowing for a more robust and transparent reading of the contributions of each determinant, both globally and locally.

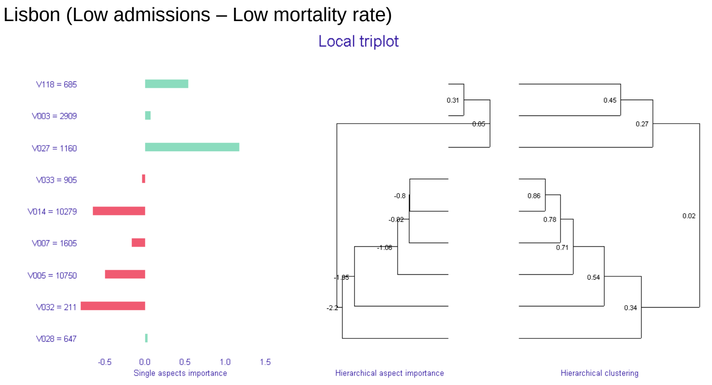
EXPLAINABLE MODEL

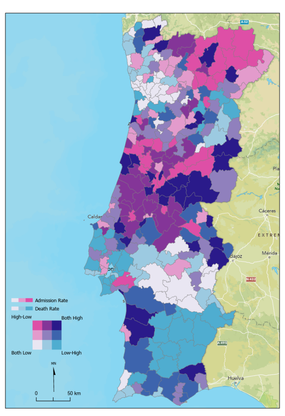
GLOBAL



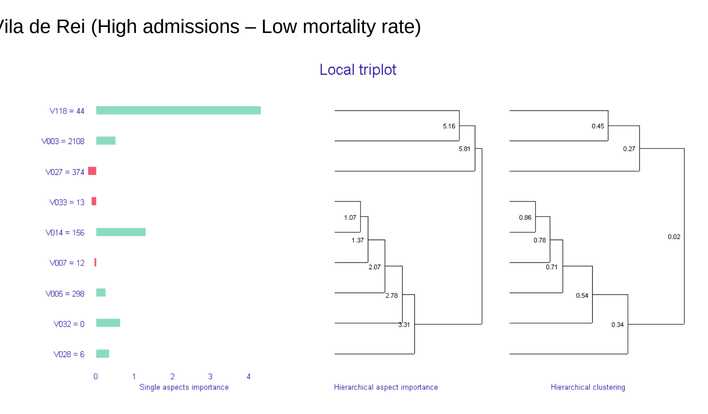


Lisbon (Low admissions – Low mortality rate)





Vila de Rei (High admissions – Low mortality rate)



LOCAL

The Matrix of Spatial Variation: Global vs. Local

The Global Average	Lisbon (Urban Focus)	Vila de Rei (Rural Focus)
V118 (Primary Sector): Universally High Risk	V118: **Mitigated Risk**	V118: **Overwhelming Risk**
V027 (Adobe walls): Universally High Risk	V027: Mitigated Risk	V027: High Risk
V014 (Pre-1919 buildings): Moderate Risk	V014: **PROTECTIVE FACTOR**	V014: **SEVERE RISK FACTOR**

V118 Resident individuals employed in the primary sector V005 Classical buildings in band V027 Buildings with adobe wall structure or loose stone masonry V014 Buildings built before 1919 V028 Buildings with another type of structure V003 Isolated classical buildings V033 Collective accommodation V007 Another type of classic building V032 Non-Classic Family Accommodations

REFERENCES

Betco, I., & Rocha, J. (2024). A relação entre o ambiente urbano e o bem-estar: Análise em Lisboa, Portugal, utilizando redes sociais. *Revista Cartográfica*, *108 SE-Artículos*, 7–28. https://doi.org/10.35424/rcarto.i108.4496

Viana, C. M., & Rocha, J. (2025). GeoAI : Entre Métodos Tradicionais e Novos Paradigmas Espaciais. In E. Reis (Ed.), *Inteligência Artificial e Geografia* (pp. 28–37). Associação Portuguesa de Geógrafos. https://www.apgeo.pt/sites/default/files/inforgeo_27_-_2025_7.pdf

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. *Ecological Indicators*, *131*, 108200. https://doi.org/https://doi.org/10.1016/j.ecolind.2021.108200

[2022.11665.BD], [UID/295/2025] https://doi.org/10.54499/UID/00295/2025