PREDICTING SOCIO-ECONOMIC AND ENVIRONMENTAL INDICATORS WITH SATELLITE IMAGES USING DEEP NEURAL NETWORKS: A CASE STUDY IN BRAZILIAN TERRITORY

Vitor Dias Souza, Samuel Vieira Ducca, Megumi Tsuru

Ph.D. Jeaneth Machicao

Prof. Dr. Pedro Luiz Pizzigatti Corrêa

Escola Politécnica da Universidade de São Paulo

megtsuru@usp.br

Objectives

This work aims to develop and adjust a model based on neural networks that can predict environmental and socioeconomic indicators within the regions of the São Francisco River Basin and Vale do Ribeira through the analysis of satellite images.

Materials and Methods

A reference paper (Jean et al, 2016) that combines satellite imagery with machine learning to predict poverty indicators in African territory was used to develop this work. The present work is divided into two study areas. During the first stage of the project, the methodology described in the article will be replicated across two important areas of the Brazilian territory: the Vale do Ribeira and the São Francisco River Basin, both of which will be run in two separate experiments, since the data granularity will be tested by census sectors in the first region and by municipalities in the second.

Two dataset types were used: the satellite images were downloaded using the Google Earth Engine API, while the indicators were obtained from two different sources. In Vale do Ribeira, socioeconomic indicators provided by the Brazilian Institute of Geography and Statistics (IBGE) were used. For the second experiment still in progress, which has been applied in the São Francisco River Basin area, socioeconomic and environmental indicators provided by the National Institute for Space Research (INPE) will be used.

Results

The adjustment of the reference paper for the Vale do Ribeira area had a satisfactory result, with a correlation coefficient between predicted and actual data of r2 = 0.37 (income indicator), as can be seen in Figure 1.

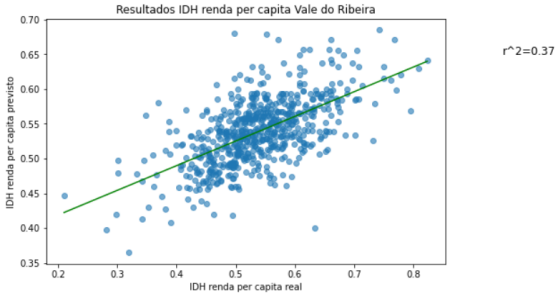

Figure 1: Comparison between the income indicator predicted by the model and the actual income in the Vale do Ribeira area

Conclusions

The replication experiment for the Vale do Ribeira area yielded promising results. It is important to note that this area has been studied by the research group in which the students are involved. A better performance is expected with the second area (São Francisco River Basin), in which the replication of the experiment will be done with a larger granularity (by municipalities) and with the increment of socioeconomic and environmental indicators datasets.

References

[1] Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. Science, 353(6301), 790–794.

[2] Triñanes, Enrico, MACHICAO, Jeaneth, & CORRÊA, Pedro. (2021). Application of a deep learning algorithm for predicting socioeconomic data through satellite images in the Vale do Ribeira [Zenodo]. https://doi.org/10.5281/zenodo.4712815