
HOUSING PRICE PREDICTION WITH COLOMBIAN REAL-STATE DATA USING MACHINE LEARNING

A PREPRINT

 **Elias Buitrago Bolivar** *

Facultad de Sistemas
Universidad ECCI
Bogota, Colombia
ebuitragob@ecci.edu.co

May 4, 2023

ABSTRACT

House price prediction is a great area of interest to the machine learning community because it provides a common framework to train and validate algorithms, techniques and methodologies, especially atkin to regression. Also, the problem of predict housing prices has gained attention from the real-state industry. Despite the overwhelming literature available on the topic in the international scientific context, Colombia lacks from studies addressing in detail the already mentioned topic with local data. One of the elements affecting this issue regard to the unavailability of specific public real-state datasets. Therefore, this paper addressed the problem of house price prediction for the Colombian context in two ways. First, a public real-state dataset obtained by web scrapping is released and documented. Second, a comparative study of performance of various well-known machine learning algorithms applied on the dataset obtained is presented. The present study tries to stablsh a baseline to increase the research about housing price prediction within the Colombian context.

Keywords housing price prediction · public datasets · machine learning · regression.

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.