



Avaliações de Imóveis Urbanos utilizando Técnicas de Mineração de Dados

Ricardo da Silva Tavares, Henrique Rego Monteiro da Hora, Renato Gomes Sobral Barcellos

O mercado imobiliário e da construção civil é um dos principais setores industriais brasileiro, este foi responsável por uma receita líquida aproximada de R\$ 2,56 bilhões, correspondendo a 4,36% do PIB brasileiro e empregando quase 2 milhões de funcionários formais. A avaliação de imóveis é uma análise complexa, cujo resultado é a determinação do valor de mercado do bem avaliado, sendo este impactado por fatores como características físicas do imóvel, geolocalização e fatores político-econômicos. Avaliações patrimoniais incorretas causam prejuízos aos cofres públicos e a iniciativa privada e as técnicas mais utilizadas atualmente, como: homogeneidade de amostras e regressão linear apresentam alguns problemas, diminuindo a precisão dessas estimativas, justificando a necessidade de aperfeiçoamento das técnicas de avaliações. Outro fator que influencia na diminuição da assertividade das avaliações patrimoniais é a dispersão de dados de amostras de mercado, isso significa não ter uma base de dados consolidada com a população de ofertas de venda de imóveis por região, impossibilitando uma visão mais ampla do mercado pelo engenheiro avaliador. Com o aumento das bases de dados de ofertas de imóveis e do poder computacional, as técnicas de Mineração de Dados tornam-se interessantes, possibilitando ao avaliador ter uma visão mais holística do mercado, permitindo-o chegar a conclusões mais assertivas. O objetivo da pesquisa é utilizar as técnicas de Mineração de Dados para aumentar a precisão das avaliações de imóveis urbanos. A metodologia utilizada engloba: (i) Estudo bibliométrico determinando os artigos mais relevantes, até março de 2020; (ii) Análise sistemática destes artigos, elencando as técnicas e resultados obtidos; (iii) Construção de um robô *crawler*, utilizando *Python*, *Beautiful Soup* e *Selenium* para *web scrapping*, possibilitando a coleta de dados dispersos nas bases de dados da *web*, nos *sites*: Zap Imóveis, Viva Real e OLX; (iv) Elaboração de *extract, transform and load* para remoção de dados inconsistentes, capazes de enviar as análises e os resultados; (v) Utilização dos algoritmos de *Machine Learning*: Regressão Linear Múltipla, *Random Forest* e *Perceptron* para cálculo das avaliações patrimoniais; (vi) Análise da assertividade das avaliações obtidas pelos algoritmos de *Machine Learning*. Como resultados espera-se a obtenção das principais variáveis que influenciam nos valores de mercado, maior precisão das estimativas de valores de mercado dos imóveis urbanos, obtidas pelos modelos de *Machine Learning* propostos, do que as técnicas mais utilizadas atualmente na Engenharia de Avaliações e determinação dos modelos de *Machine Learning* mais assertivos.