

Precisão na avaliação automatizada de imóveis: o potencial da técnica de Random Forest

Arquimedes S. F. Macedo¹, Pedro Augusto²

UFRGS/IME, Porto Alegre, RS

Manuela de Castro³

UFRGS/IME, Porto Alegre, RS

A avaliação de imóveis é uma tarefa crucial para diversos setores da economia, como o mercado imobiliário, o setor bancário e o setor de seguros. No entanto, obter uma estimativa de valor para um imóvel não é uma tarefa trivial. Além de que, essa atividade desafiadora, pode afetar negativamente os consumidores finais que necessitam de uma base para guiar sua compra ou venda de imóvel. Para a administração pública não é diferente.

A precisão na avaliação de imóveis é de suma importância para garantir a justiça na aplicação de tributos sobre bens imóveis, desta maneira, estudos como [1] e [2] alertam, que, avaliações de baixa qualidade, podem levar a uma distribuição desigual da carga tributária entre os contribuintes, violando o princípio da isonomia. Além disso, para [1], o método comparativo de avaliação, recomendado pela NBR 14653-2 [3], é frequentemente aplicado de maneira inadequada pelos municípios, resultando em distorções adicionais de valores.

Uma das principais causas dessa dificuldade é a utilização de modelos de regressão linear, uma técnica estatística simples que se torna limitada em problemas com alta dimensionalidade [4]. Essa modelagem, por preços hedônicos, teorizada por [5], tem sido amplamente utilizada na avaliação de imóveis, sendo definida como um conjunto de preços observados com as quantidades específicas de características associadas com cada bem. Seu objetivo é determinar o preço de um produto a partir da análise de suas características individuais [5]. No caso de um imóvel, por exemplo, é possível modelar uma função preço que possui como variáveis de entrada características como localização geográfica, idade do prédio, índices de criminalidade do bairro, etc., além da conjuntura econômica do país, a oferta e demanda por imóveis na região, entre outros.

Recentemente, técnicas de machine learning têm sido aplicadas na avaliação de imóveis, permitindo a criação de modelos mais precisos e eficientes do que os métodos tradicionais baseados em regressão linear. Neste contexto, o modelo de Random Forest tem se destacado como uma das técnicas mais promissoras para a precificação imobiliária [6].

Esta técnica de aprendizado de máquina consiste em um conjunto de árvores de decisão independentes e diversificadas que são combinadas para produzir uma única previsão [7]. Utilizamos esse modelo para avaliar o preço de imóveis em um conjunto de dados com mais de 1,1 milhões de observações, considerando 30 variáveis, tais como quantidade de quartos, suítes, área, posição geográfica, estatísticas criminais e populacionais, acessibilidade, tempo de atividade do anúncio, entre outras.

A partir da utilização da linguagem Python e da biblioteca scikit-learn, obtivemos um modelo com uma precisão média de 89% na previsão do preço dos imóveis. A partir deste modelo foram geradas inferências para pontos na região estudada (Centro de Porto Alegre, RS).

¹arquimedes.macedo@ufrgs.br

²00335558@ufrgs.br

³Orientadora de Iniciação Científica. Email: manuela.castro@ufrgs.br

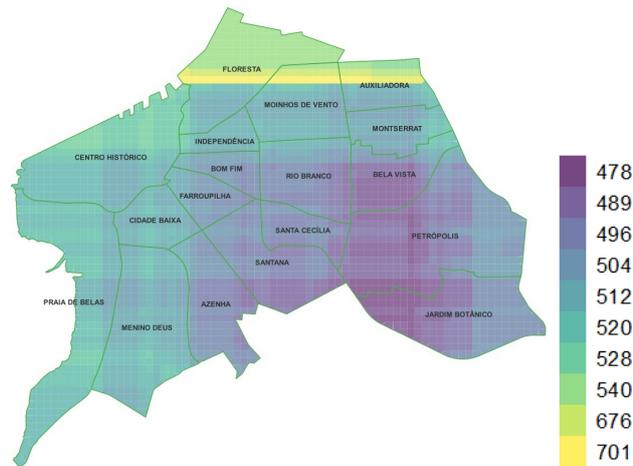


Figura 1: Variação geográfica do preço inferido, em milhares de reais.

Desta forma, a aplicação do modelo de Random Forest na avaliação de imóveis apresentou resultados promissores, podendo ser, portanto, uma alternativa mais eficiente e precisa em comparação com modelos tradicionais baseados em regressão linear, oferecendo uma solução viável e confiável para a avaliação de imóveis.

Agradecimentos

Agradecemos à NIVU (startup especializada em análise de dados imobiliários) pela generosa disponibilização dos dados de anúncios atuais, que foram utilizados nesta pesquisa.

Referências

- [1] C. M. de Cesare e E. M. P. Cunha. **Avaliação em Massa de Imóveis para Fins Fiscais: discussão, análise e identificação de soluções para problemas e casos práticos**. Brasília: Ministério das Cidades, 2012. ISBN: 9788579580284.
- [2] S. Poeta. “Consequências da Defasagem da Planta de Valores Imobiliários de IPTU da Cidade de Porto Alegre/RS”. Tese de doutorado. UFRGS, 2019. DOI: hdl.handle.net/10183/198032.
- [3] **NBR 14653-2. Avaliação de bens, imóveis urbanos**. Associação Brasileira de Normas Técnicas. Rio de Janeiro, 2004.
- [4] J. Neter, W. Wasserman e M. H. Kutner. **Applied linear statistical models: Regression, analysis of variance, and experimental designs**. R. D. Irwin, 1974.
- [5] S. Rosen. “Hedonic Prices and Implicit Markets: product differentiation in pure competition”. Em: **Journal of political economy** 82.1 (1974), pp. 34–55. DOI: [10.1086/260169](https://doi.org/10.1086/260169).
- [6] J. Hong, H. Choi e W.- Sung Kim. “A House Price Valuation Based on the Random Forest Approach: the mass appraisal of residential property in South Korea”. Em: **International Journal of Strategic Property Management** 24.3 (2020), pp. 140–152. DOI: [10.3846/ijspm.2020.11544](https://doi.org/10.3846/ijspm.2020.11544).
- [7] T. Hastie, R. Tibishirani e J. Friedman. **The Elements of Statistical Learning: data mining, inference and prediction**. Springer, 2016.