



Universidade Federal De Pernambuco
Centro de Informática
Graduação em Ciência da Computação

**Classificação e Predição de Preços de Imóveis a Partir de Dados
Estruturados e Não Estruturados**

Proposta de Trabalho de Graduação

Aluno: **Matheus Barbosa Domingues Feliciano**
Orientador: **Luciano Barbosa**

Recife
Agosto, 2018

Resumo

O crescimento do uso da internet como o principal meio de consulta e geração de conteúdo, nos últimos anos, refletiu numa quantidade imensa de dados na web. Esses dados podem ser utilizados de diversas formas, como análises estatísticas, modelagem de dados e previsões. O principal objetivo desse trabalho é predizer, a partir de dados estruturados e não-estruturados de imóveis, o valor que deveriam ter. O mercado imobiliário vem em crescente no Brasil, como uma forma de investimento indo de encontro com a crise, comparar qual o impacto da utilização e não utilização de dados não-estruturados para a predição pode melhorar a precisão de definição de valores desses imóveis.

Sumário

Contexto	2
Objetivo	3
Cronograma	4
Possíveis Avaliadores	5
Referências	6
Assinaturas	7

Contexto

A avaliação imobiliária, que é o processo de estimar o preço de imóveis, é crucial para ambos compradores e vendedores como uma base para a negociação e transação. O mercado imobiliário tem um papel vital em todos os aspectos da nossa sociedade contemporânea [1]. Com o crescimento da geração de dados online sobre imóveis, o caminho mais intuitivo tomado para a análise desses foi o de estimar esse preço através dos atributos desses imóveis, ou seja, os dados já estruturados, como quantidade de quartos, localização, área interna, etc.

Porém, outra forma de talvez aumentar a acurácia e precisão de um algoritmo que faça essa estimativa seria utilizar em conjunto com os dados estruturados, dados não-estruturados, neste caso, as descrições dos imóveis, comentários sobre ele e avaliações de usuários. Todos esses são exemplos de dados não-estruturados em forma de texto. Que seriam tratados de forma a inferir características adicionais aos imóveis que poderiam influenciar no seu preço, e dessa forma aumentando a precisão do algoritmo.

A análise de dados não-estruturados, mais especificamente de textos, para a predição de preços de produtos é uma tendência que vem aumentando nos últimos anos. Uma competição promovida pela *Kaggle* e a *Mercari* [2], neste ano, teve como foco a construção de um algoritmo que sugeriria automaticamente o preço correto de produtos [2], sendo fornecidos dados não-estruturados em forma de textos de descrições escritos por usuários com informações como marca, categoria do produto e condições.

Utilizando essas duas formas de avaliação imobiliária através de algoritmos de análise de dados estruturados e não-estruturados, poderemos comparar a utilização das duas formas de avaliação e como um interfere na precisão e acurácia do outro quando utilizadas em conjunto. Desse modo pretende-se criar uma ferramenta que consiga utilizar os dois métodos separadamente, assim como também em conjunto e compará-los.

Objetivo

O objetivo deste trabalho é prever preços de imóveis de acordo com dados obtidos da web, devidamente pré-processados, e comparar resultados somente com dados estruturados e com dados estruturados e não-estruturados. Com o intuito de definir o impacto dos dados não-estruturados, neste caso, textos de descrição do imóvel, na classificação e previsão do preço desses imóveis.

Cronograma

Atividades	Set	Out	Nov	Dez
Preparação dos dados				
Estudo da Literatura				
Implementação dos algoritmos				
Análise dos resultados				
Apresentação				

Possíveis Avaliadores

Ricardo Bastos Prudêncio

Referências

- [1] - YOU, Quanzeng et al. Image-Based Appraisal of Real Estate Properties. IEEE Transactions on Multimedia, [S.l.], 27 jun. 2017. Volume: 19, Problema: 12, p. 2751 - 2759. Disponível em: <<https://ieeexplore.ieee.org/document/7937942/>>. Acesso em: 29 ago. 2018.
- [2] - MERCARI Price Suggestion Challenge: Can you automatically suggest product prices to online sellers?. 2018. Disponível em: <<https://www.kaggle.com/c/mercari-price-suggestion-challenge>>. Acesso em: 29 ago. 2018.

Assinaturas

Aluno: **Matheus Barbosa Domingues Feliciano**

Orientador: **Luciano Barbosa**