

### III Workshop de Dissertações de Mestrado do PPgSI (2016)

#### FICHA DA PESQUISA<sup>i</sup>

DADOS GERAIS				
<b>Título do projeto de pesquisa</b>	Análise preditiva de valores de imóveis na Grande São Paulo com base em notícias sobre economia e criminalidade			
<b>Orientando</b>	Fabio Purcino Aragão			
<b>Orientador(es)</b>	Norton Trevisan Roman			
<b>Momento atual</b>	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input checked="" type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre
<b>Qualificação</b>	<input type="checkbox"/> Qualificação já realizada em: ___/___/___ <input checked="" type="checkbox"/> Qualificação planejada para: 25/08/2016			
<b>Defesa</b>	Prazo máximo para depósito: 26/02/2018		Depósito planejado para: 15/01/2017	
<b>Linha e Área de pesquisa</b>	<input type="checkbox"/> Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input type="checkbox"/> Eng. de Software <input type="checkbox"/> IHC		<input checked="" type="checkbox"/> Inteligência de Sistemas: <input type="checkbox"/> IA <input checked="" type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico	
<b>Área de aplicação</b>	<input type="checkbox"/> Ambientes Corporativos <input type="checkbox"/> Educação <input type="checkbox"/> Bioinformática <input type="checkbox"/> Educação a Distância <input type="checkbox"/> Biometria <input type="checkbox"/> Internet <input type="checkbox"/> Economia <input type="checkbox"/> Jogos	<input checked="" type="checkbox"/> Linguagem Natural <input type="checkbox"/> Redes Sociais <input type="checkbox"/> Linguística <input type="checkbox"/> Robótica <input type="checkbox"/> Processos de Negócio <input type="checkbox"/> Saúde <input type="checkbox"/> Química <input type="checkbox"/> [outro – escrever]		
<b>Publicações</b>	[EM ELABORAÇÃO] Artigo de Revisão Sistemática a ser submetido à Revista Brasileira de Sistemas de Informação - ISYS			

DESCRIÇÃO DO PROJETO DE PESQUISA	
<b>Contextualização / motivação</b>	O projeto é aplicado à análise preditiva de preços do mercado imobiliário específico da Grande São Paulo. Além dos indicadores econômicos há também fatores de bem-estar social que influenciam na lei da oferta e da procura.
<b>Problema de pesquisa</b>	Observa-se hoje no mercado imobiliário uma variação muito alta de preços em função de indicadores econômicos e da especulação, distorcendo a relação entre o valor do imóvel e o benefício que ele proporciona ao morador que busca fatores de bem-estar social, muitas vezes externados por meio da internet.
<b>Objetivo geral da pesquisa</b>	O objetivo do trabalho é elaborar um modelo de análise preditiva dos preços de imóveis na Grande São Paulo com base em notícias sobre a economia e criminalidade em cada região.
<b>Trabalhos relacionados</b>	Em uma esfera global Fávero(2005), Carmo(2014) utilizaram o conceito de hedonismo definido pelo conjunto de fatores de bem-estar procurados pelas pessoas na hora de investir em moradia elaborando modelos de regressão linear com base em dados disponibilizados pelas prefeituras e outros órgãos estatísticos, nem sempre refletindo os reais desejos do consumidor. Por sua vez Quintanilha(2012) analisa especificamente os efeitos da criminalidade sobre os preços do mercado de habitação em São Paulo de forma empírica usando dados da polícia militar.
<b>Justificativa e relevância</b>	Esta pesquisa é justificada pela necessidade de controle e planejamento urbano com base não na especulação mas no preço potencial que uma região pode ter como reflexo de estímulos na economia e iniciativas que melhorem a segurança local. A relevância do projeto reside no fato de não estar baseado em indicadores enviesados mas buscar os fatos ocorridos na mídia, um meio que exerce influência direta e indireta sobre o consumidor permitindo a simulação dos preços futuros em função do reflexo de políticas econômicas e de segurança implantadas pela iniciativa pública e privada.
<b>Proposta para Solução</b>	Será adotada uma nova abordagem com base em um modelo de regressão linear onde seus coeficientes serão estimados em função da máxima verossimilhança de uma amostra de preços passados com a amostra resultante de algoritmos de classificação com a quantidade de notícias sobre os tipos de crime e polaridade econômica.
<b>Dados</b>	Foi criado um corpus com 8911 textos integrais de notícias sobre criminalidade e 2242 textos integrais sobre economia do site do jornal “Estado de São Paulo” abrangendo o período de 2007 a 2014. Também serão utilizados os preços reais do metro quadrado residencial publicados pelo SECOVI-SP.
<b>Validação</b>	O modelo resultante será submetido a avaliações cruzadas com os preços reais para análise do erro médio e da correlação de Pearson.
<b>Limitações</b>	O projeto está limitado aos indicadores de criminalidade e economia enviesados à opinião de um único veículo. Não foram abordados outros fatores de bem-estar como proximidade de escolas, facilidade de transporte e áreas de lazer. A crise econômica brasileira também pode afetar a validade dos experimentos.

<b>Resultados esperados</b>	Contribuições científicas: O corpus e os algoritmos resultantes serão públicos e contribuem na área de pesquisa sobre processamento de língua natural e na área de ciências econômicas mensurando a influência da mídia sobre os preços praticados no mercado imobiliário.
	Contribuições tecnológicas: Será desenvolvido um software de apoio que classificará as notícias e irá calcular os coeficientes. Este software pode ser usado também para classificar posts de redes sociais, blogs e outros tipos de textos em língua natural.

<b>MÉTODOS DE PESQUISA</b> [para cada item, marque todas as opções adequadas para seu projeto]				
<b>Gênero</b>	<input type="checkbox"/> Pesquisa teórica	<input type="checkbox"/> Pesquisa prática	<input checked="" type="checkbox"/> Pesquisa empírica	<input type="checkbox"/> Pesquisa metodológica
<b>Natureza</b>	<input type="checkbox"/> Pesquisa básica/pura	<input checked="" type="checkbox"/> Pesquisa aplicada		
<b>Objetivo</b>	<input type="checkbox"/> Pesquisa descritiva	<input checked="" type="checkbox"/> Pesquisa exploratória	<input type="checkbox"/> Pesquisa explicativa	<input type="checkbox"/> [outro – escrever]
<b>Abordagem</b>	<input checked="" type="checkbox"/> Pesquisa quantitativa	<input type="checkbox"/> Pesquisa qualitativa	<input type="checkbox"/> Pesquisa mista (quali-quant)	
<b>Procedimento(s) técnico(s)</b>	<input checked="" type="checkbox"/> Pesquisa experimental <input type="checkbox"/> Pesquisa bibliográfica <input type="checkbox"/> Pesquisa documental <input type="checkbox"/> Pesquisa <i>ex-post-facto</i> <input type="checkbox"/> Pesquisa de levantamento	<input type="checkbox"/> Pesquisa com <i>survey</i> <input type="checkbox"/> Estudo de caso <input type="checkbox"/> Pesquisa participante <input type="checkbox"/> Pesquisa-ação <input type="checkbox"/> Pesquisa etnográfica	<input type="checkbox"/> Pesquisa netnográfica <input type="checkbox"/> Teoria fundamentada em dados ( <i>grounded theory</i> ) <input type="checkbox"/> Ciência do projeto ( <i>Design science research</i> ) <input type="checkbox"/> [outro – escrever]	
<b>Fonte(s) de dados</b>	<input type="checkbox"/> pesquisa de laboratório	<input checked="" type="checkbox"/> pesquisa de campo	<input type="checkbox"/> pesquisa bibliográfica	<input type="checkbox"/> [outro – escrever]
<b>Técnica(s) / Instrumento(s) de coleta de dados</b>	<input type="checkbox"/> medição <input type="checkbox"/> questionário <input type="checkbox"/> entrevista <input type="checkbox"/> grupos focais	<input type="checkbox"/> formulário <input type="checkbox"/> <i>benchmark</i>	<input type="checkbox"/> observação (direta / participante) <input type="checkbox"/> diário de campo / notas de campo <input checked="" type="checkbox"/> análise documental (ou de artefatos) <input type="checkbox"/> [outro – escrever]	
<b>Técnica(s) de análise de dados</b>	Análise quantitativa: <input type="checkbox"/> Estatística descritiva <input checked="" type="checkbox"/> Estatística inferencial <input type="checkbox"/> [outro – escrever]	Análise qualitativa: <input type="checkbox"/> Análise de conteúdo <input type="checkbox"/> Análise do discurso <input type="checkbox"/> [outro – escrever]		

<b>CRONOGRAMA</b> [altere conforme necessário]																																					
	2015												2016												2017												
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	
Estudo da literatura			X	X	X																																
Revisão sistemática						X	X																														
Definição da proposta								X																													
Execução da proposta									X	X	X																										
Coleta/obtenção de dados									X	X	X																										
Análise dos dados												X	X																								
Validação da proposta												X	X	X																							
Preparo da qualificação												X	X	X	X																						
Exame de qualificação																				X																	
Escrita da dissertação																					X	X	X	X													
Escrita de artigo (RS)										X	X	X	X	X	X	X	X	X	X																		
Depósito da dissertação																																			X		
(Outros)																																					

<sup>i</sup> Esta ficha é uma adaptação da usada no “VIII Workshop de Teses e Dissertações em Sistemas de Informação (WTDSI 2015)” realizado como parte do “XI Simpósio Brasileiro de Sistemas de Informação (SBSI 2015)”