



4º SEMINÁRIO NACIONAL DESAPROPRIAÇÃO E REASSENTAMENTO

2025



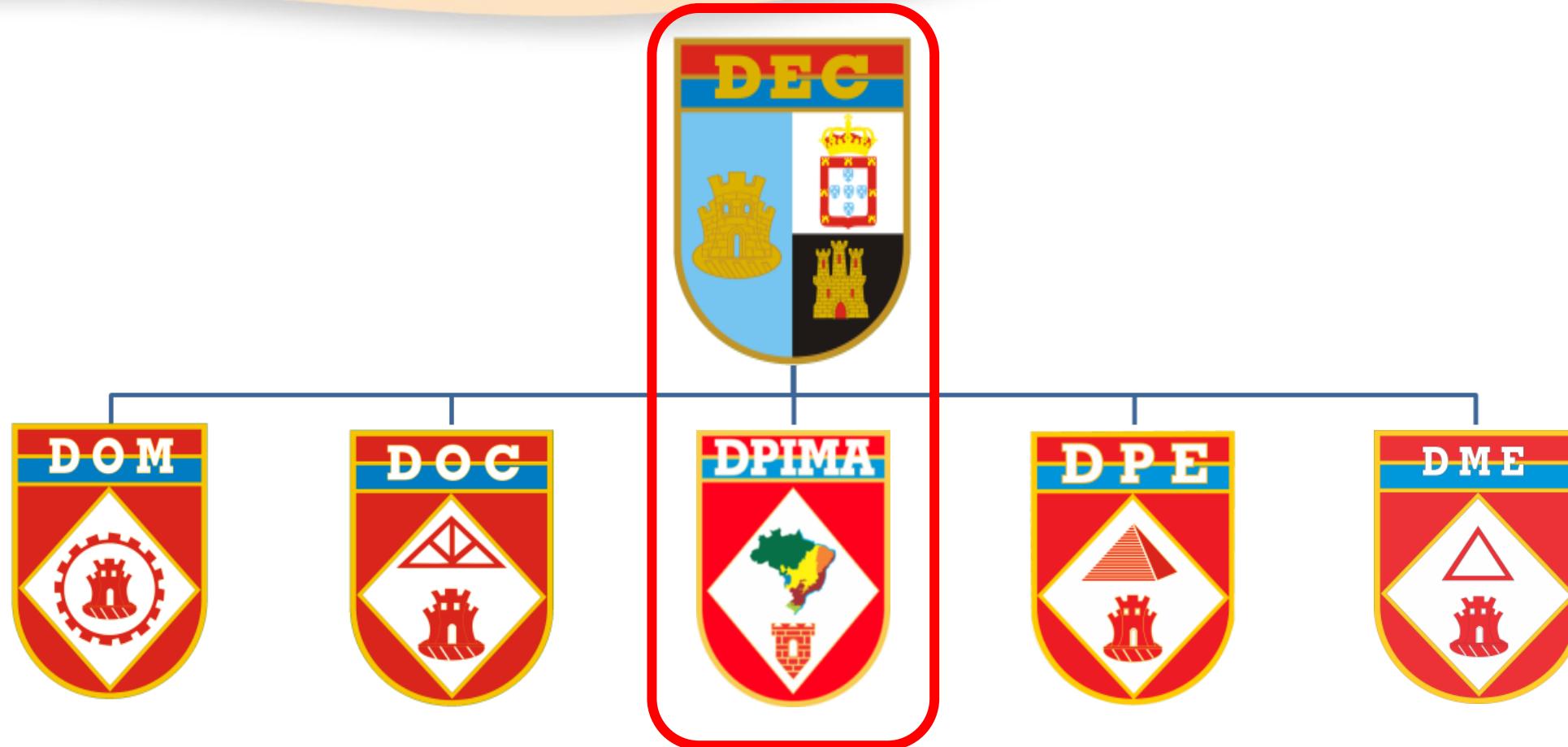
MINISTÉRIO DOS
TRANSPORTES



AVALIAÇÃO DE IMÓVEIS URBANOS E RURAIS COM USO DE APRENDIZADO DE MÁQUINA

Estudos de Caso

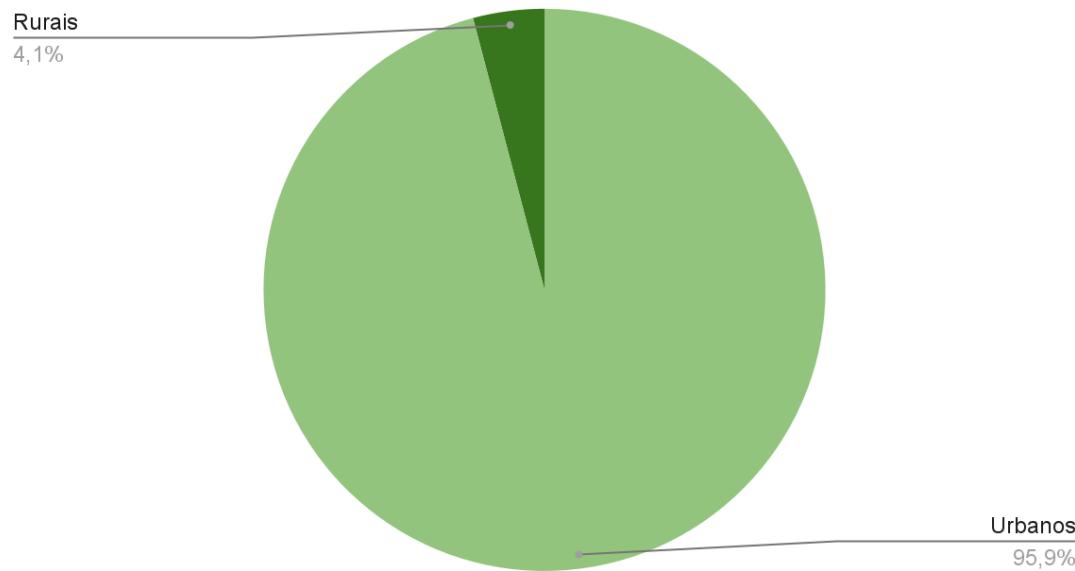
CONTEXTO



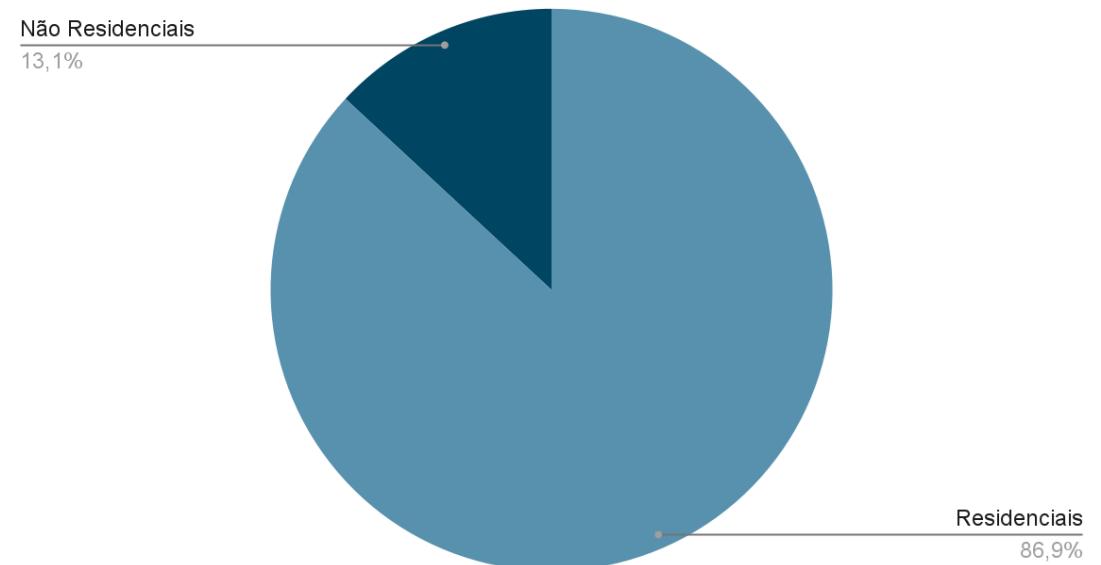
CONTEXTO

- O EB possui **mais de 20 mil** parcelas imobiliárias da União sob sua administração. A União, **mais de 700 mil**.

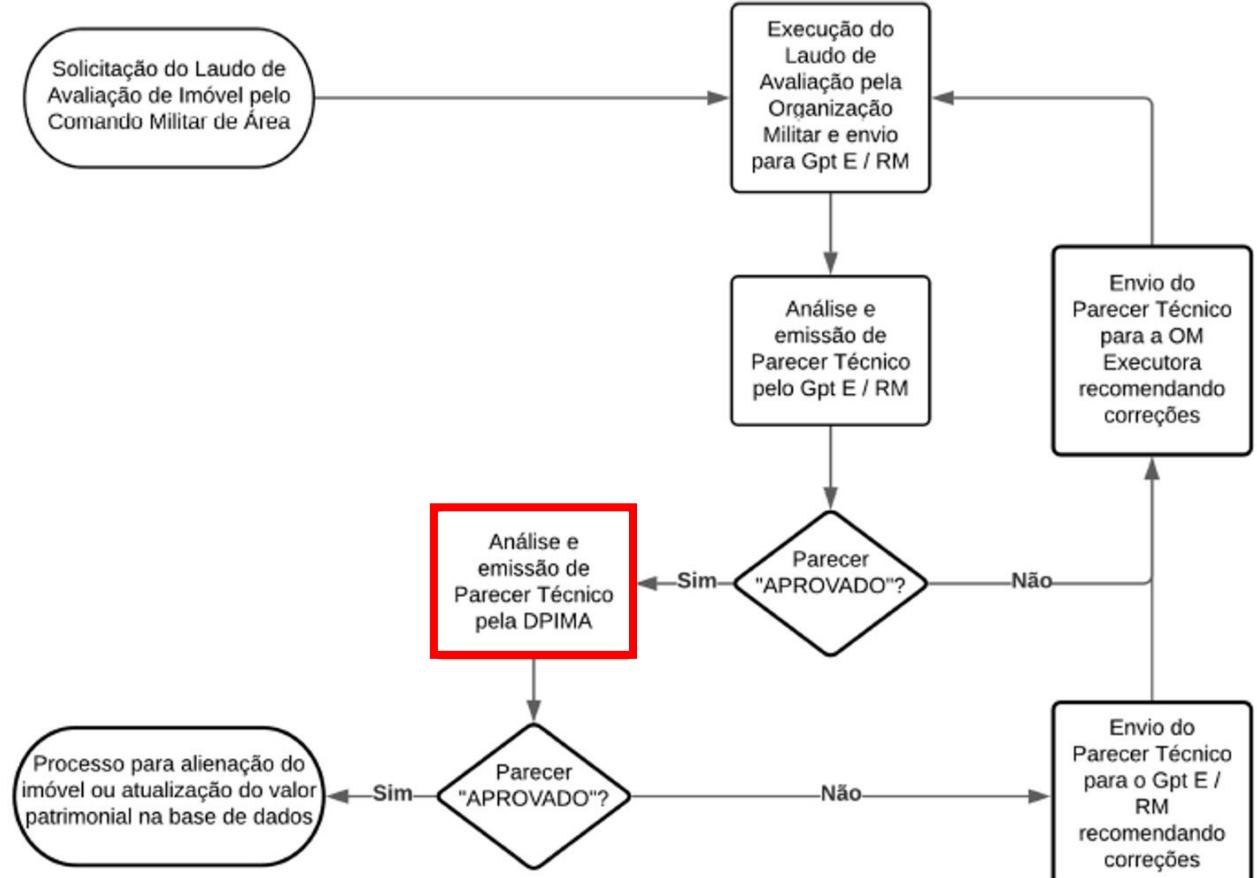
População (aprox. 21500 unidades imobiliárias)



População (aprox. 21500 unidades imobiliárias)



CONTEXTO



CONTEXTO

- Como propor modelos capazes de construir **estimativas de valor** dos imóveis administrados pelo EB com razoável aptidão à generalização?

Trabalho	R ² > 57%	Abrangência nível País	Modelagem linear	Modelagem não linear	Ajuste de hiperparâmetros	Componentes espaciais	Mais de uma vocação de imóvel	Interpretabilidade
Dantas et al. 2010 [3]	✓	✓	✓			✓		✓
Park e Bae 2015 [4]	✓		✓	✓	✓			
Kiely e Bastian 2020 [1]			✓	✓	✓	✓	✓	
Hagenauer e Helbich 2022 [2]		✓	✓	✓		✓		
Tchuente e Nyawa 2022 [5]	✓	✓	✓	✓	✓	✓		
Solução proposta	✓	✓	✓	✓	✓	✓	✓	✓

BASES DE DADOS



Fonte: EB /
SPU / IBGE /
GitHub Prof.
Dr. Bernardo
Furtado.

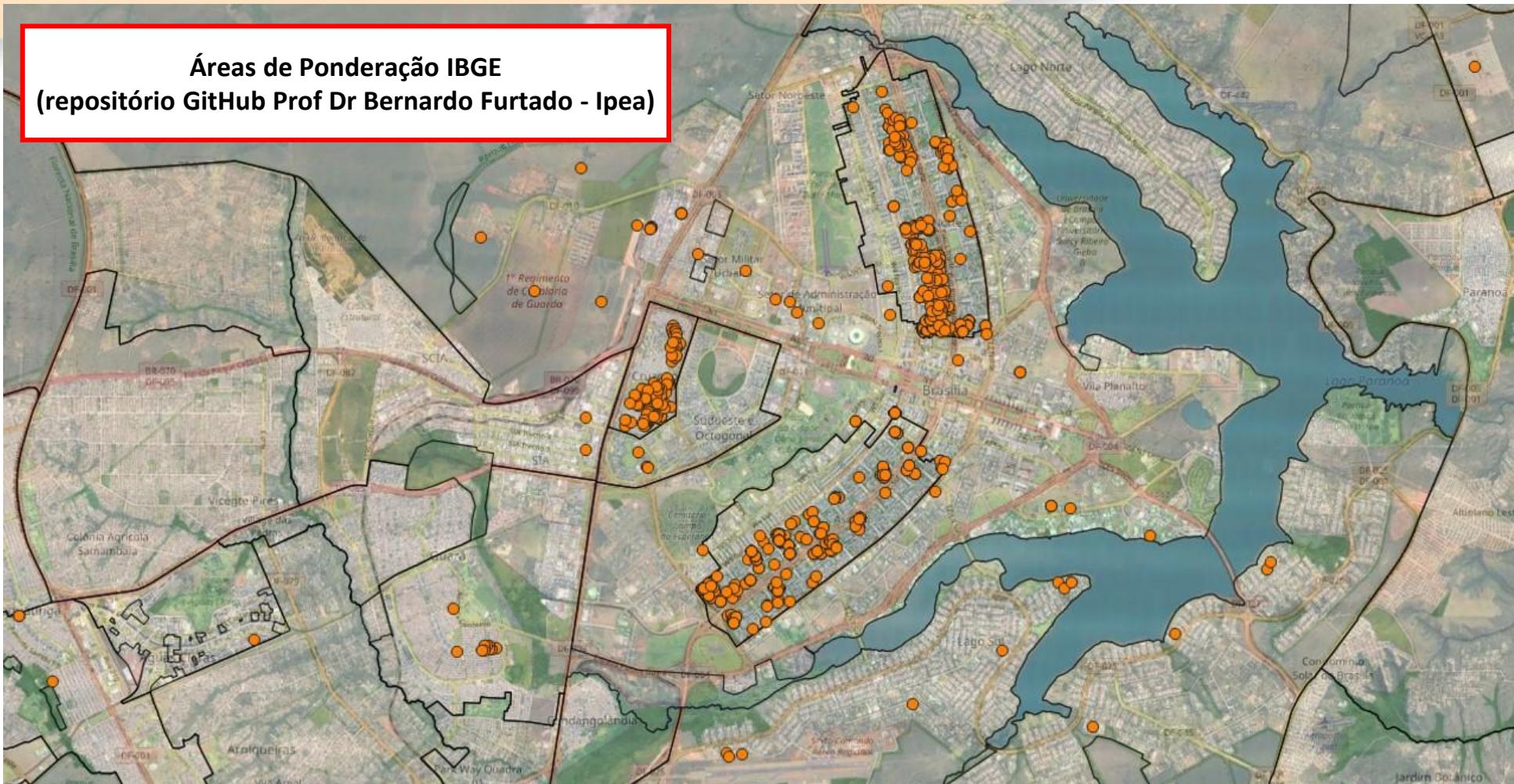
Fonte	Qtde de instâncias urbanas com valor conhecido	Participação relativa
EB	258	5,6%
SPU	4340	94,4%
EB e SPU	4598	100,0%

METODOLOGIA



- **Pré-processamento:** geocodificação, união dos arquivos gerados (incluindo elementos geográficos IBGE e Google), **codificação binária** de variáveis categóricas nominais.
- **Análise exploratória:** verificação gráfica e seleção de atributos (com SEM).
- Aplicação de **testes de normalidade** e de **transformações matemáticas** sobre os atributos.
- Validação cruzada: **10 folds**.
- **Testes intradomínio:** treinamento e teste em base **EB/SPU (Cenário A)**.
- **Testes intradomínio:** treinamento em base **EB/SPU** e teste em **base EB (Cenário B)**.
- **Comparação** dos modelos entre si (lineares e não lineares).
- Métricas de avaliação utilizadas: **coeficiente de determinação (R^2)** e **raiz do erro quadrático médio (RMSE)**.
- Interpretabilidade com **valor de Shapley** (Teoria dos Jogos).

EXPERIMENTOS REALIZADOS



EXPERIMENTOS REALIZADOS



IMÓVEIS URBANOS

data **zap+**

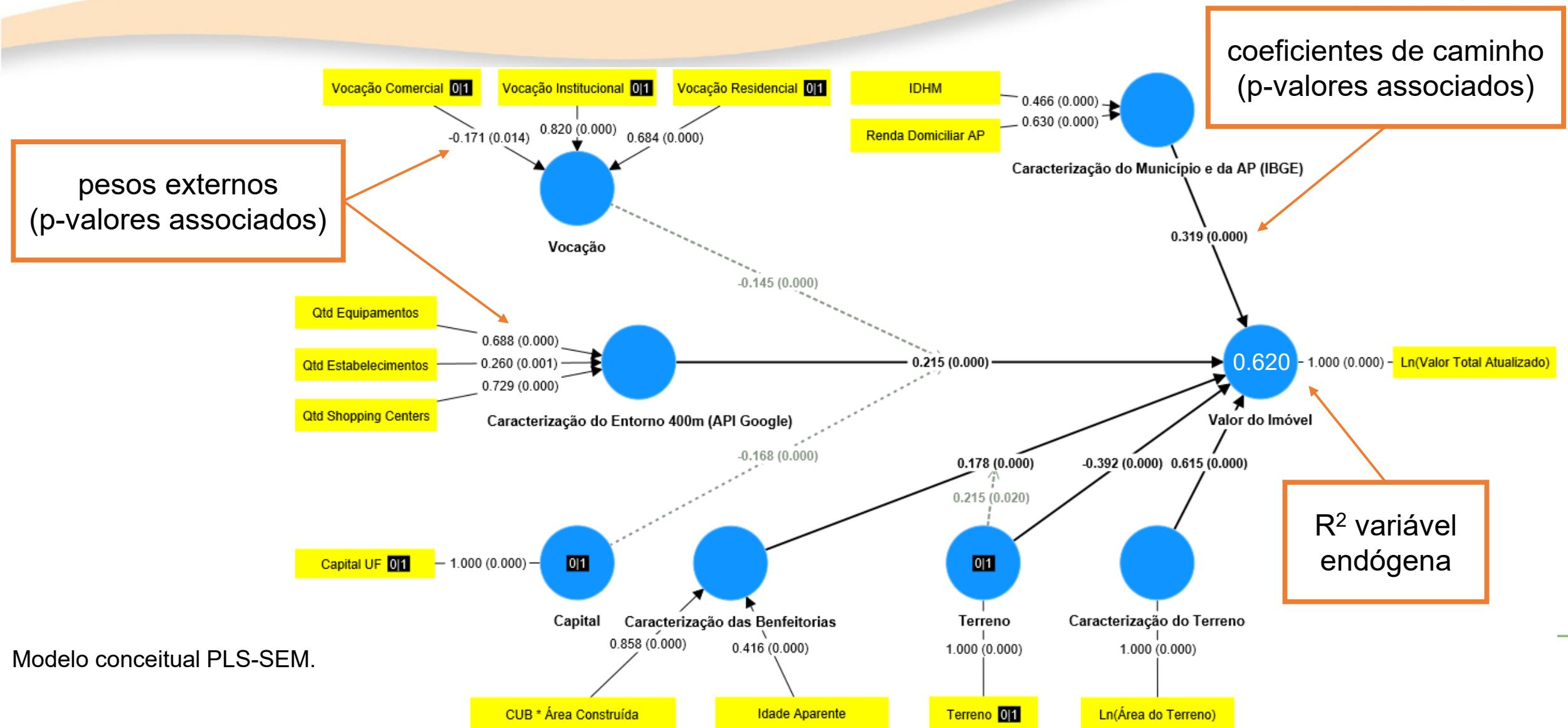


Google Places API

Pontos de interesse
em raio de 400 m de cada imóvel:

- agregado equipamentos
- atrações turísticas
- delegacias
- escolas
- hospitais
- parques
- universidades
- aeroportos
- estações de metrô
- estações de trem
- estações de VLT
- paradas de ônibus
- agregado estabelecimentos
- cafeteria
- lojas
- padarias
- restaurantes
- **shopping centers**
- supermercados

EXPERIMENTOS REALIZADOS

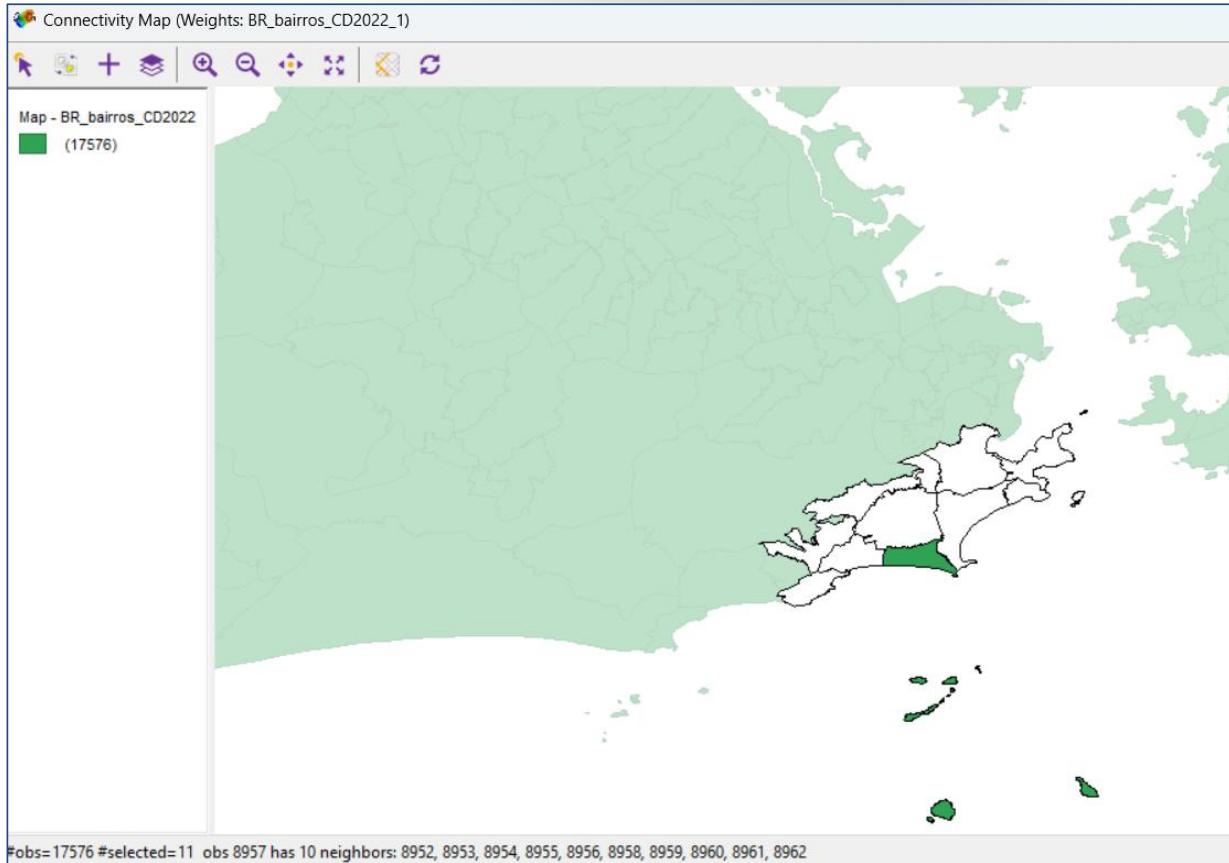


EXPERIMENTOS REALIZADOS

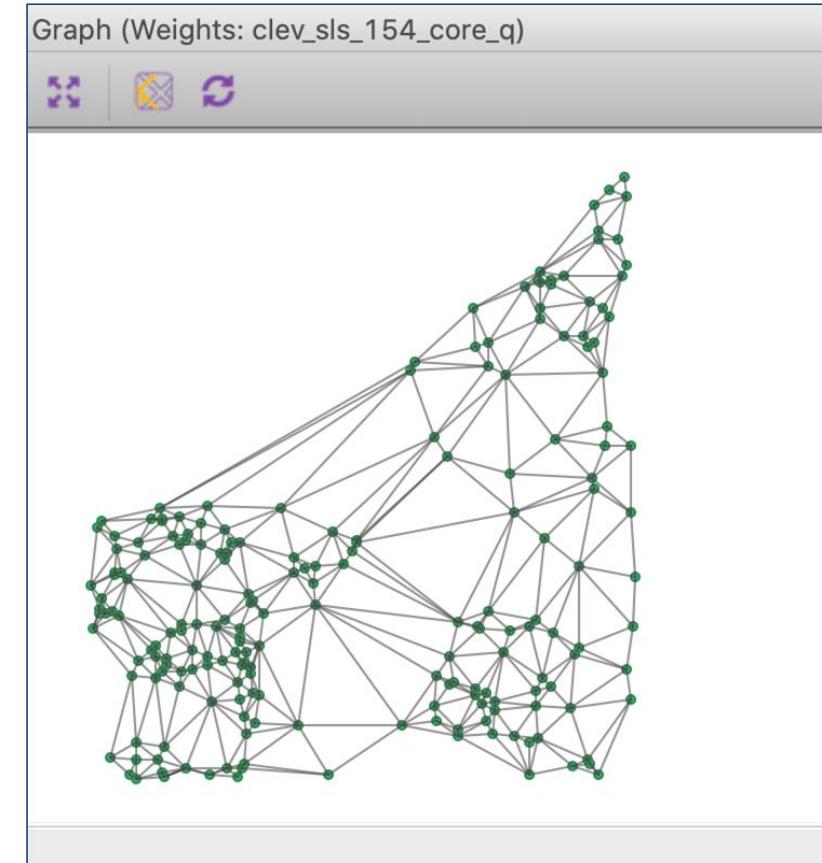
- Atributos **urbanos** efetivamente utilizados em cada nível de modelagem:

Variável	Nível de Modelagem
<i>Capital UF</i>	U_1, U_2, U_3
<i>Vocação do Imóvel</i>	U_1, U_2, U_3
<i>Índice de Desenvolvimento Humano Municipal</i>	U_1, U_2, U_3
<i>Ln(Área do Terreno)</i>	U_1, U_2, U_3
<i>Custo Unitário Básico * Área Construída</i>	U_1, U_2, U_3
<i>Idade Aparente</i>	U_2, U_3
<i>Renda Domiciliar AP</i>	U_3
<i>Pontos de Interesse API Google Places</i>	U_3
<i>Ln(Valor Total Atualizado)</i>	U_1, U_2, U_3

EXPERIMENTOS REALIZADOS



Mapa de conectividade construído para a região do Rio de Janeiro.



Grafo utilizado para construção de matrizes de vizinhança.

EXPERIMENTOS REALIZADOS



IMÓVEIS URBANOS

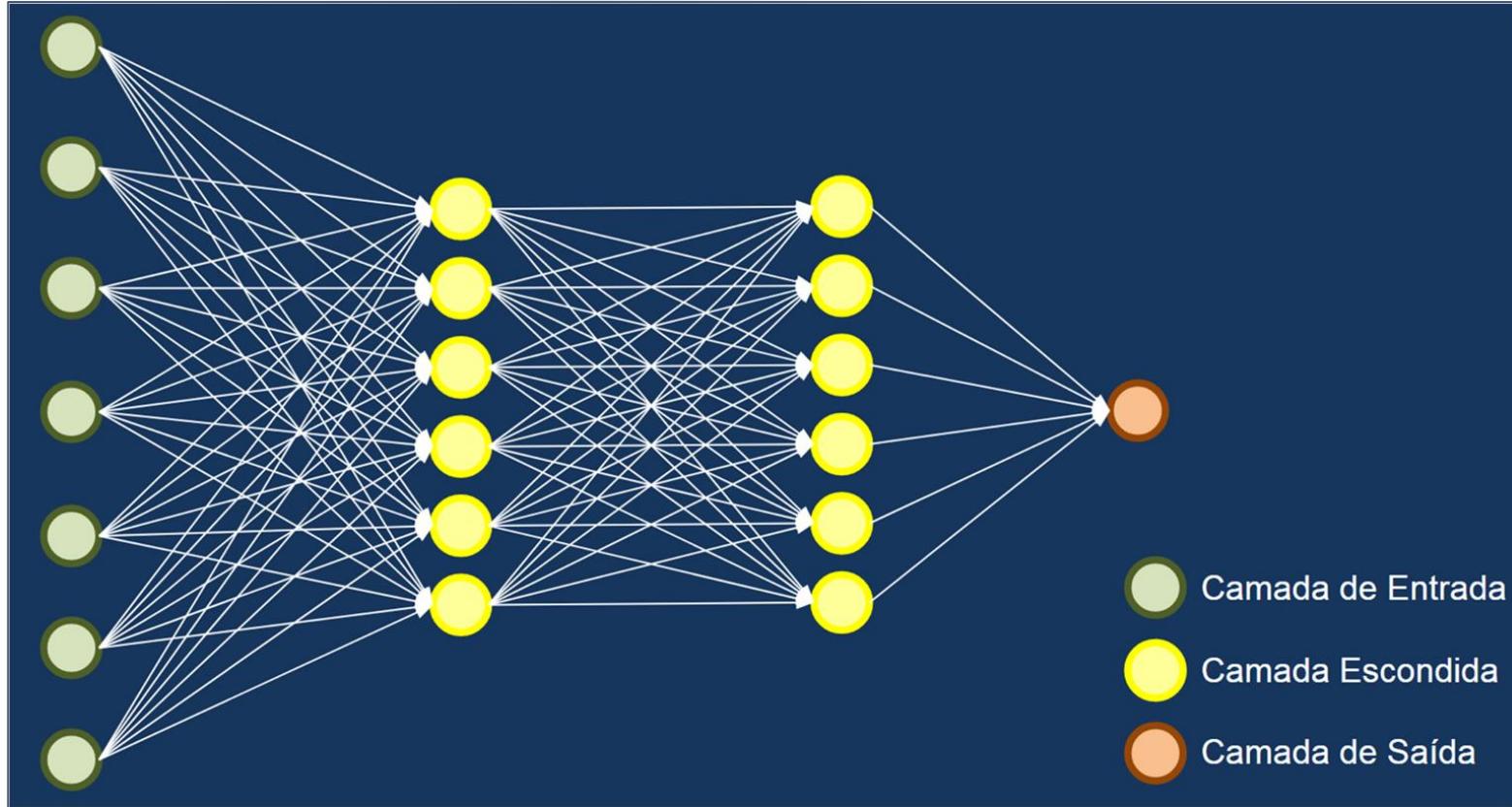
Parâmetro	<i>SGDRegressor</i> linear	<i>ANN MLPRegressor</i> não linear	<i>XGBRegressor</i> não linear
Fração de Treinamento	53,3%	53,3%	53,3%
Fração de Validação	13,3%	13,3	13,3%
Fração de Teste	33,3%	33,3	33,3%
Tipo de taxa de aprendizagem	constante	constante	constante
Valor da taxa de aprendizagem	0,0001	0,01	0,3
Função de custo	erro quadrático	erro quadrático	erro quadrático
Termo de regularização	<i>L2</i>	<i>L2</i> (multiplicador $\alpha = 0,1$)	<i>L1</i> ($\alpha = 5$) e <i>L2</i> ($\lambda = 1$)
Arquitetura básica	não se aplica	(5, 5) ^a em U ₁ , (6, 6) em U ₂ e (8, 8) em U ₃	<i>max_depth</i> = 5

^a camadas ocultas da rede neural artificial



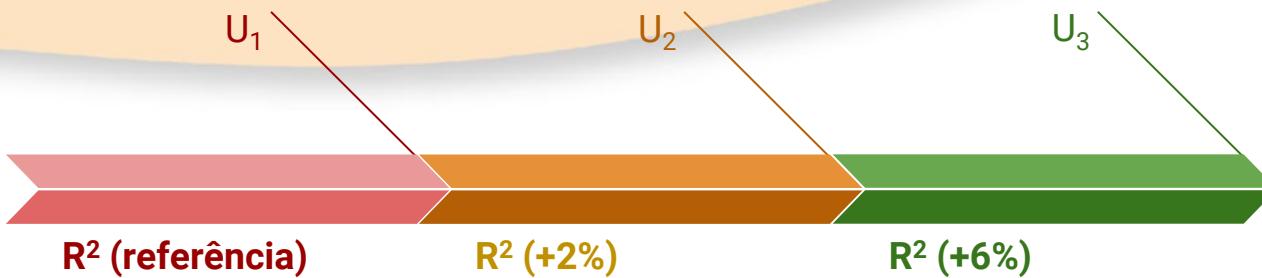
Treinamento - Validação - Teste

EXPERIMENTOS REALIZADOS



Arquitetura de rede neural artificial (RNA) construída pela DPIMA.

RESULTADOS OBTIDOS

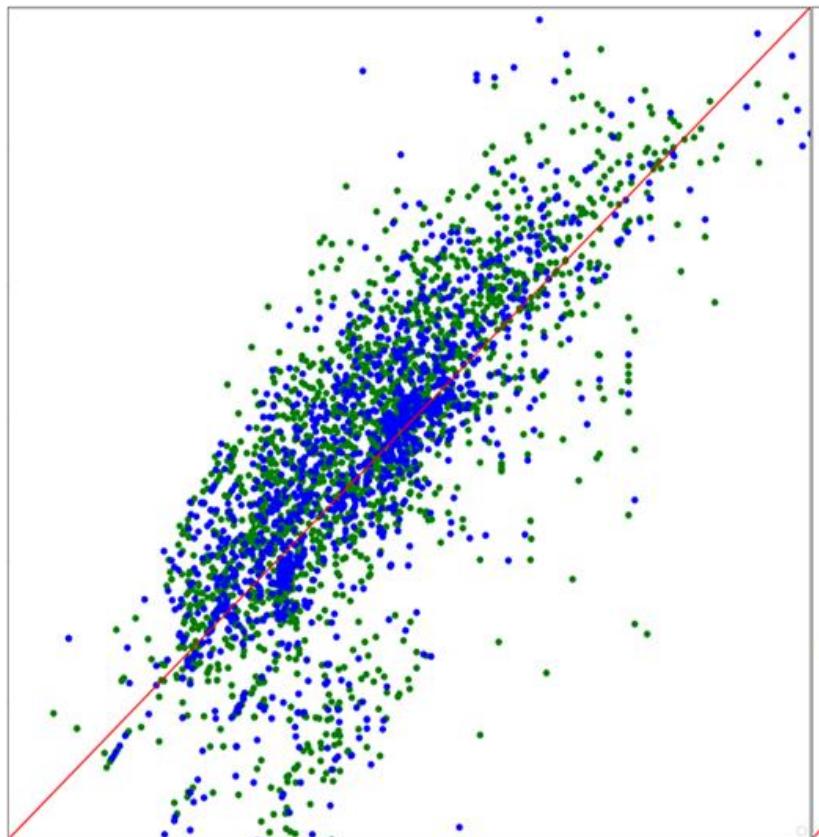


Modelos U_3 (Cenário)	R^2 médio <i>10 folds</i>	dp R^2 <i>10 folds</i>	R^2 <i>train</i>	R^2 <i>test</i>	$RMSE$ <i>train</i>	$RMSE$ <i>test</i>
OLS (A)	NA ^a	NA ^a	61,9%	52,3%	1,28	1,73
Regressão Espacial (A)	NA ^a	NA ^a	63,4%	55,7%	1,26	1,39
<i>SGDRegressor</i> linear (A)	61,2%	0,047	61,9%	61,1%	1,27	1,31
<i>SGDRegressor</i> linear (B)	60,7%	0,044	61,7%	50,3%	1,29	1,35
ANN não linear (A)	73,8%	0,048	74,8%	69,3%	1,04	1,17
ANN não linear (B)	71,8%	0,041	73,8%	71,3%	1,06	1,02
<i>XGBRegressor</i> não linear (A)	86,2%	0,019	95,7%	83,6%	0,43	0,85
<i>XGBRegressor</i> não linear (B)	86,7%	0,035	94,8%	80,6%	0,47	0,84

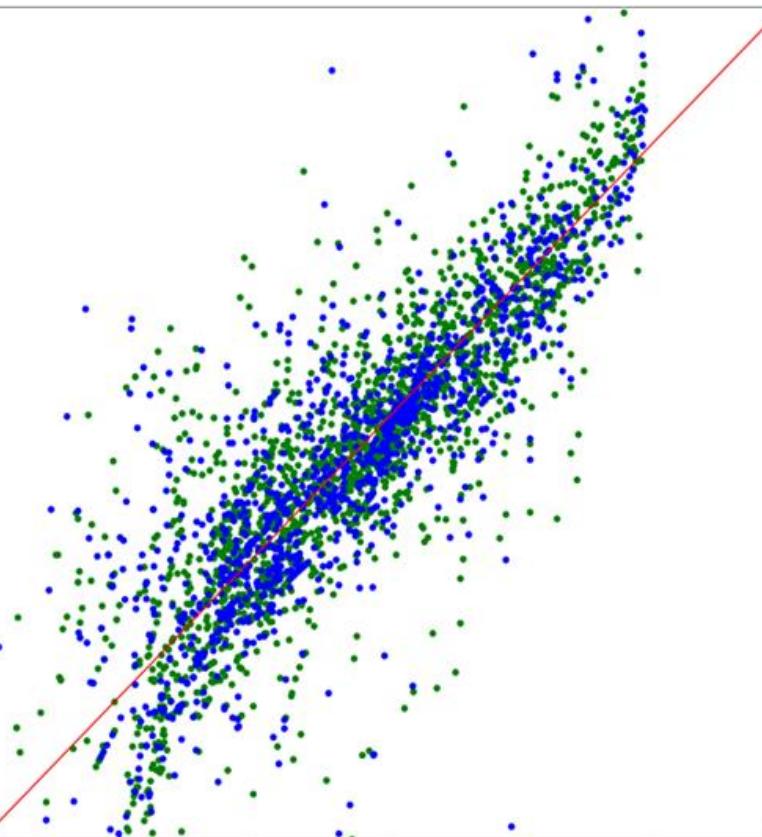
^a não se aplica

RESULTADOS OBTIDOS

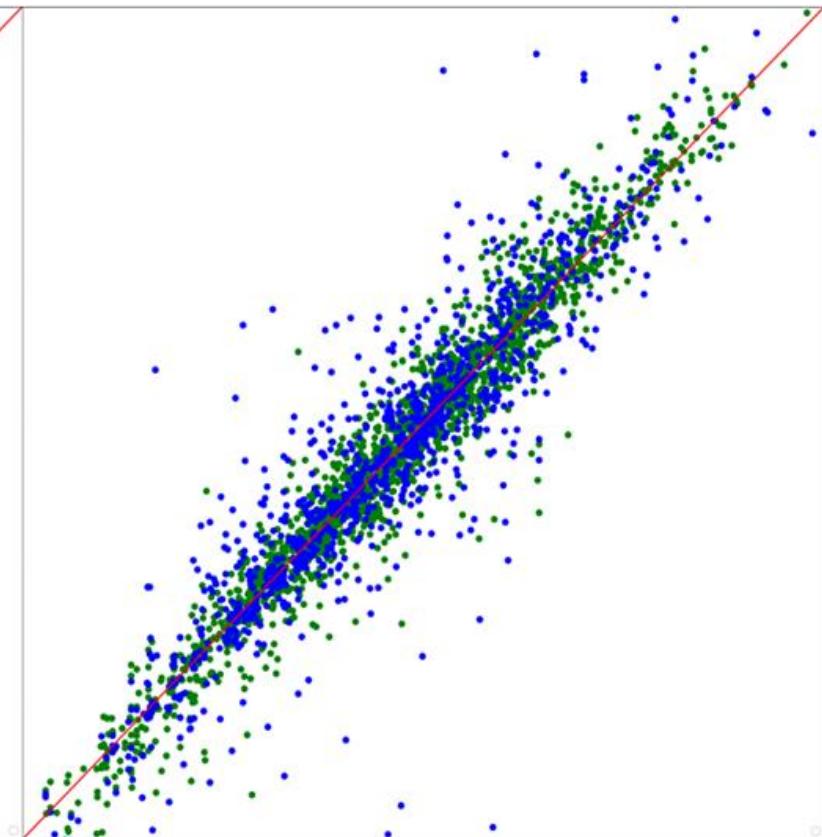
U_3 *SGDRegressor* (Cenário A)



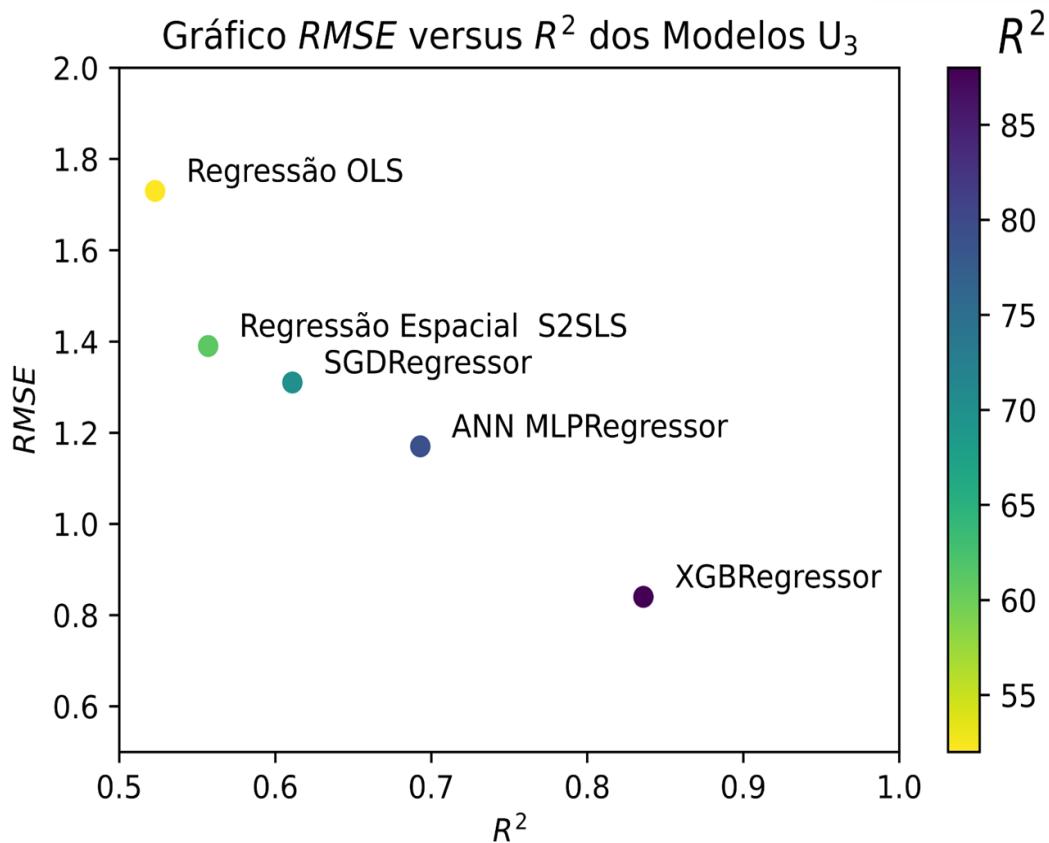
U_3 *ANN MLPRegressor* (Cenário A)



U_3 *XGBRegressor* (Cenário A)



RESULTADOS OBTIDOS

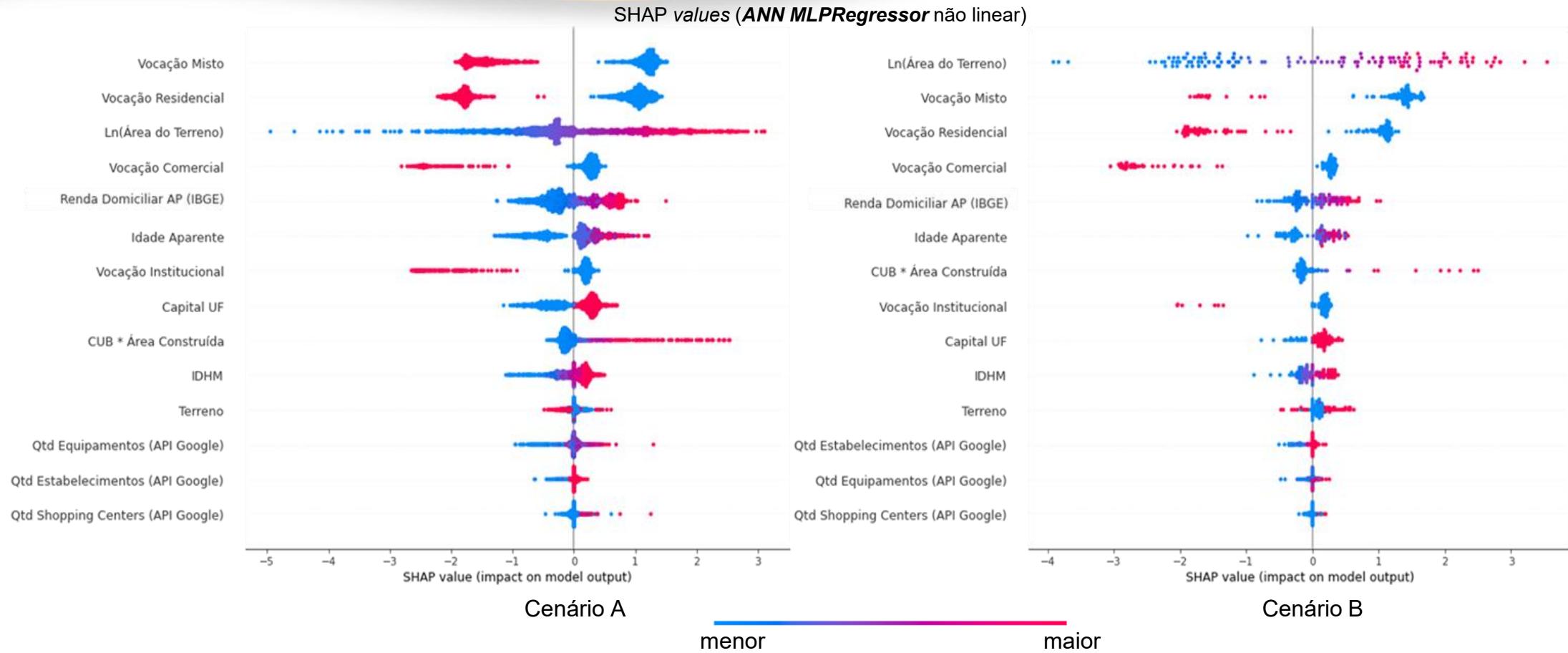


R^2

55 60 65 70 75 80 85

Modelo	Limite IC Inferior	Limite IC Superior	Amplitude IC
ANN MPLRegressor U ₁	-51,0%	+109,0%	160,0%
ANN MPLRegressor U ₂	-50,9%	+107,9%	158,8%
ANN MPLRegressor U ₃	-51,1%	+98,8%	149,9%
XGBRegressor U ₁	-43,9%	+79,8%	123,7%
XGBRegressor U ₂	-43,8%	+79,6%	123,4%
XGBRegressor U ₃	-41,9%	+73,3%	115,2%

RESULTADOS OBTIDOS



RESULTADOS OBTIDOS



MAIO 2022



JULHO 2022

$$+ 2x$$

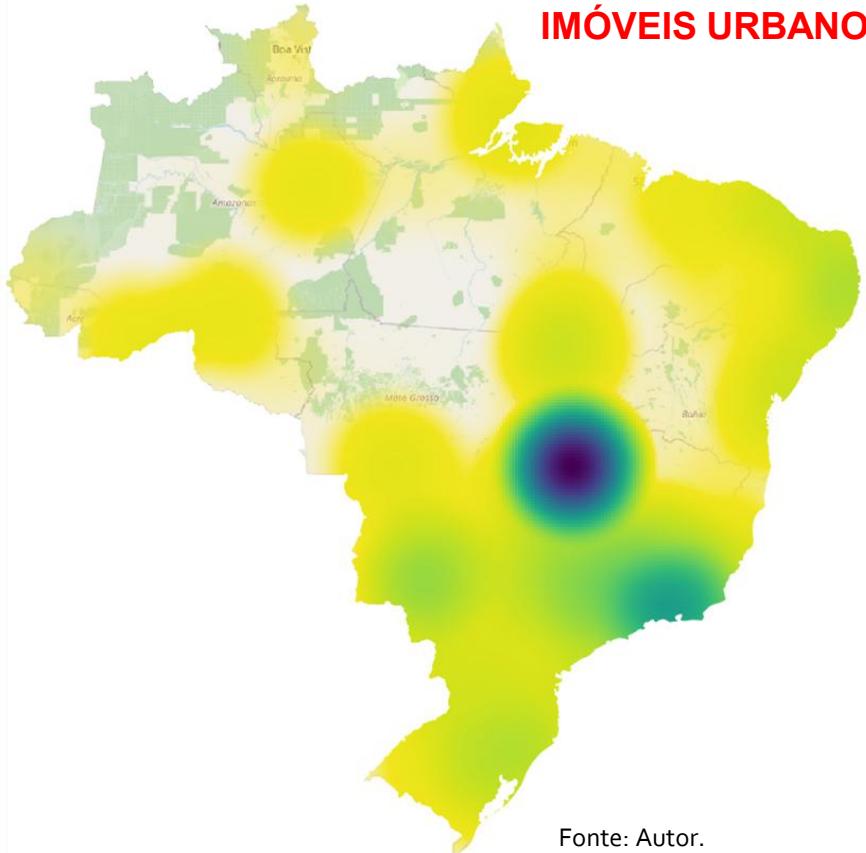


MINISTÉRIO DOS
TRANSPORTES



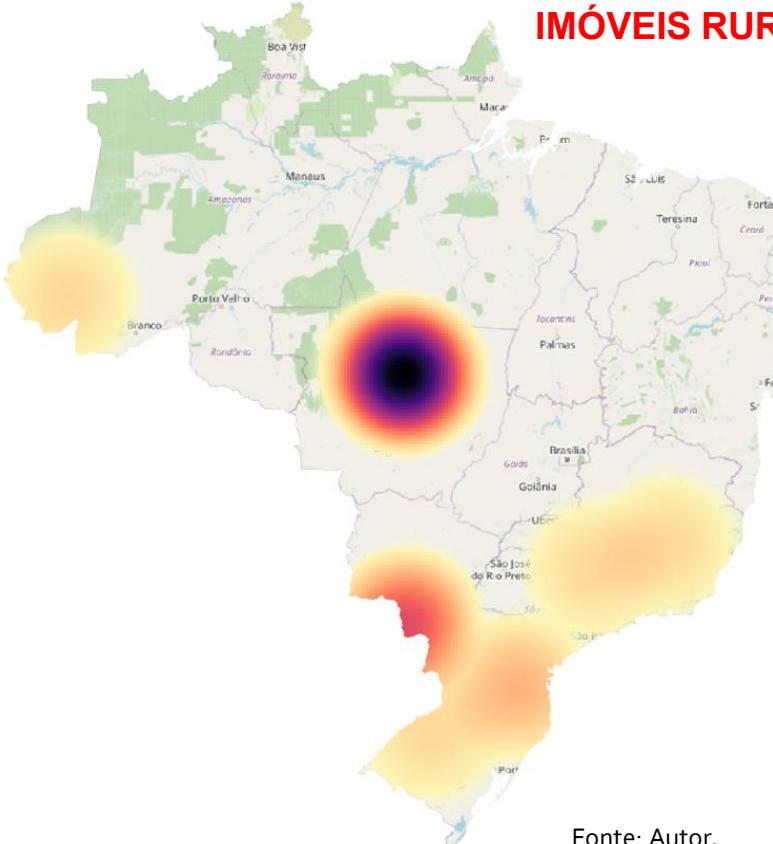
CONSIDERAÇÕES FINAIS

IMÓVEIS URBANOS



Fonte: Autor.

IMÓVEIS RURAIS



Fonte: Autor.

OBRIGADO!