



# XXXI Congresso de Iniciação Científica da UNESP

## Predizendo PIB das cidades paulistas utilizando imagens noturnas de satélite e Machine Learning

João Pedro Donaire Albino, Prof. Dr. Clayton Pereira  
Câmpus de Bauru, Faculdade de Ciências



### Introdução

A pobreza é um problema dificilmente controlado e trabalhado por métodos efetivos de combate. Arelado a isso também temos uma escassez de dados relevantes que possam mensurar a mesma. Além disso, o mapeamento sobre indicadores econômicos sempre foi algo custoso no Brasil. Por exemplo, o orçamento do Censo de 2010 pelo Instituto Brasileiro de Geografia e Estatística (IBGE) fora calculado em R\$ 1,677 bilhão (Fonte: IBGE).

Outrossim trazer métodos tecnológicos poderia auxiliar em grande parte nessa demanda.

### Objetivos

O projeto teve como escopo principal criar um sistema totalmente automatizado para extrair, estruturar, carregar e classificar imagens de satélites noturnos do Estado de São Paulo, sendo todos os arquivos de dados abertos. Dessa forma, a principal pergunta da pesquisa era: É possível correlacionar intensidade luminosa com indicadores econômicos utilizando dados abertos e Machine Learning?

### Material e Métodos

As linguagens utilizadas dentro processo foram Python e R. Além disso, durante a pesquisa foram aplicados conceitos de classificação KMeans e Support Vector Machine (SVM).

Os padrões de classificação se basearam em dados do IBGE e conceituações dos indicadores do Produto Interno Bruto (PIB). Ademais, para a obtenção das imagens dos mapas, foram utilizados arquivos da *National Center for Environmental Information* (NOAA) divididos em imagens de 10 Km<sub>2</sub>.

Por fim, para catalogar todas as imagens obtidas com as suas respectivas cidades correspondentes no estado paulista foi utilizada a API do Google Maps.

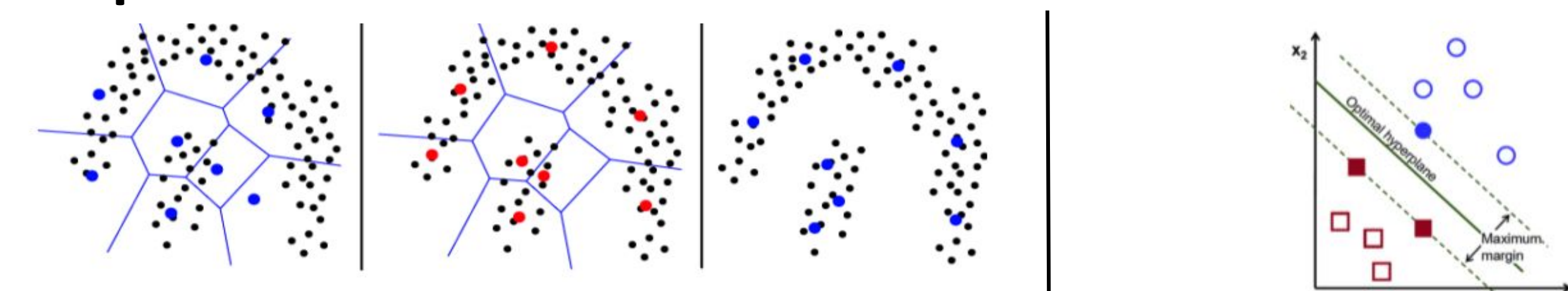


Figura 1: Ilustrações, respectivamente, dos métodos KMeans e SVM

### Financiamento



### Resultados

A pesquisa resultou em um sistema totalmente integrado desde a obtenção dos dados até a classificação. Toda a cadeia foi desenvolvida em Python.



Figura 2: Automação do sistema

Ao todo foram obtidas 3.075 imagens em conjunto com 65% das cidades do estado catalogadas com a API do Google Maps. Os formatos armazenados desses arquivos foram TIF e PGM, sendo eles posteriormente convertidos em arquivos CSV.

Ademais, a classificação utilizada para alcançarmos os resultados foi a avaliação utilizada pelo World Bank (Fonte: *New country classifications income level*), sendo obtidas acurácias de 45.81% com KMeans e 52.22% com SVM.

		SVM		
		Baixo	Médio	Alto
PIB	Baixo	643	573	13
	Médio	84	168	7
	Alto	8	67	11

Figura 3: Resultados obtidos

### Discussão

Para um início de estudos, a pesquisa se mostrou promissora tanto pela automação quanto pelos resultados obtidos.

Entretanto, existem pontos importantes para um futuro desenvolvimento, como por exemplo, uma comparação com outras técnicas de classificação supervisionadas e não-supervisionadas.

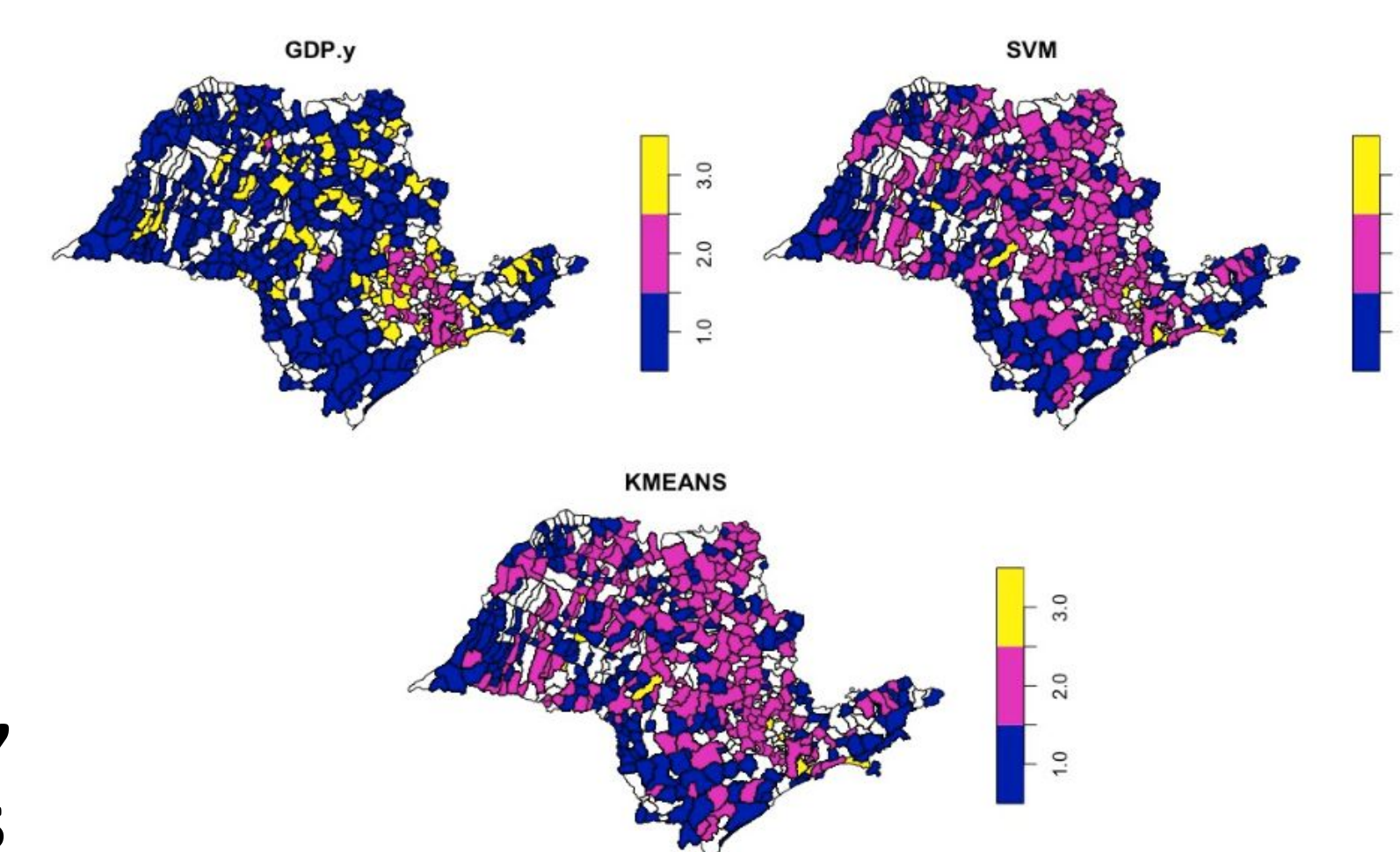


Figura 4: Classificadores

### Conclusões

Por fim, objetivo principal era responder se havia ou não correlação entre as intensidades luminosas com indicadores econômicos. Os resultados, por mais que iniciais, mostraram-se instigantes fornecendo alternativas às avaliações socioeconômicas.

### Bibliografia

- 1 G. Alejandro. Me. po.: o que te. n. li.. 2004.
- 2 Barros R.; Henriques R.; Mendonça R.. A est. ina.: Desigualdade e pob. no br.. 2001, 29.
- 3 Ebener, S.; Et al. We. to hea.: Mod. the dis. of in. per capita at the sub. lev. us. nig. lig. ima. 2005.
- 4 Jean, N.; Et al. Tran. Lea. from Deep Fea. for Rem. Sen. and Pove. Map. 2015.