

## Predizendo classes de PIB no Estado de São Paulo utilizando imagens de satélites noturnos e Machine Learning

João Pedro Donaire Albino, Prof. Dr. Clayton Pereira, Câmpus de Bauru, Faculdade de Ciências, Sistemas de Informação, joaopdalbino@gmail.com

Palavras Chave: *machine learning, mapas noturnos, predição.*

### 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. Entretanto, há alguns anos, vem sendo trabalhada a relação entre intensidades luminosas e indicadores socioeconômicos.

### Objetivo

O projeto tem como escopo principal criar um sistema totalmente automatizado para extrair, estruturar, carregar e classificar imagens de satélites noturnos do Estado de São Paulo, sendo estes de origem open-source.

### Materiais e Métodos

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

Os padrões de classificação se basearam em dados open-source do Instituto Brasileiro de Geografia (IBGE) e Estatística e conceituações indicadores do Produto Interno Bruto (PIB).

Ademais, para a obtenção das imagens dos mapas, foram utilizados dados abertos da *National Center for Environmental Information* (NOAA) divididos por imagens de 10 Km<sup>2</sup>. Fora possível encontrar cerca de 65% das cidades do Estado de São Paulo somente utilizando o *Google Api Maps*.

### Resultados e Discussão

A pesquisa procurou obter dados de forma estruturada e automatizada, propondo **maneiras mais intuitivas para a análise e classificação dos dados**.

Figura 1. Automação escrita em Python



Por fim, foi possível obter dados relevantes quando classificados pelos dois métodos. Utilizando PIB e dados do Banco Mundial como referência, foram

obtidas acurácias de **45.81%** com KMeans e **52.22%** com SVM.

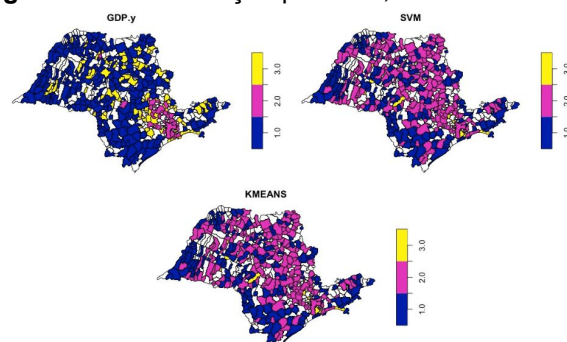
Figura 2. Matriz de Confusão

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

### Conclusões

A pesquisa se mostrou muito 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 3. A classificação pelo PIB, SVM e KMeans



### Agradecimentos

Agradeço ao professores Clayton Pereira e João Paulo Papa pelas orientações e auxílios na realização da pesquisa.

G. Alejandro. *Me. po.: o que te. n. li.*. 2004.

Barros R.; Henriques R.; Mendonça R.. *A est. ina.: Desigualdade e pob. no br.*. 2001, 29.

Ebener, S.; Et al. *We. to hea.: Mod. the dis. of in. per capita at the sub. lev. us. nig. lig. ima.*. 2005.

Jean, N.; Et al. *Tran. Lea. from Deep Fea. for Rem. Sen. and Pove. Map.*. 2015.