Identificação de desmatamento na Amazônia em imagens de satélite

Cassiano Zago (DEPEN)
Pedro J. Fernandes (UFF)
Washington R. Dias (Prefeitura de Vitória – ES)

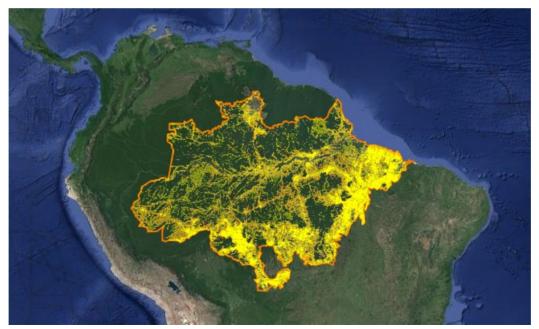
https://github.com/pjfernandes/amazon_deforestation_bootcamp_enap

Introdução

Desmatamento na Amazônia

Mais de 20% da floresta foi desmatada durante os últimos 30 anos (KHANNA et al., 2017)

KHANNA, Jaya et al. Regional dry-season climate changes due to three decades of Amazonian deforestation. Nature Climate Change, v. 7, n. 3, p. 200-204, 2017.



http://terrabrasilis.dpi.inpe.br/

Introdução

Impactos ambientais causados pelo desmatamento



Erosão http://www.agencia.cnptia.embrapa.br/



Redução da qualidade da água



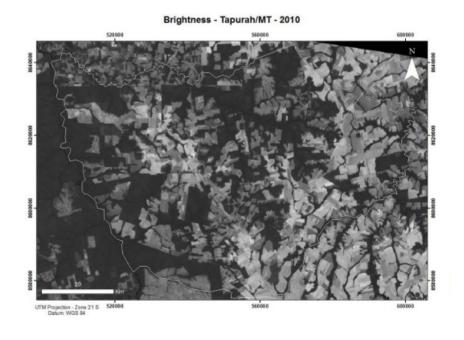
Alterações no clima local e regional

https://redesustentabilidade.org.br

Introdução (contexto)

Monitoramento do desmatamento por imagens de Sensoriamento Remoto

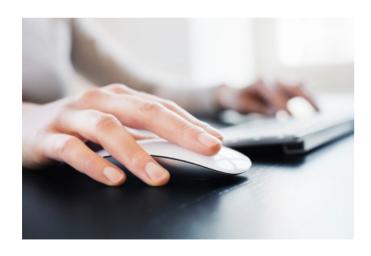


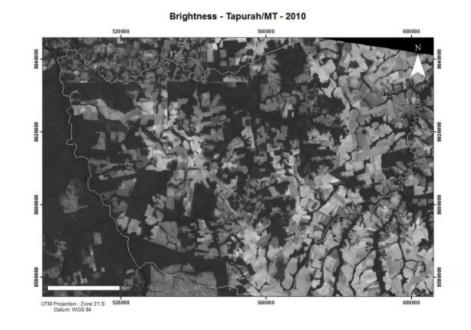


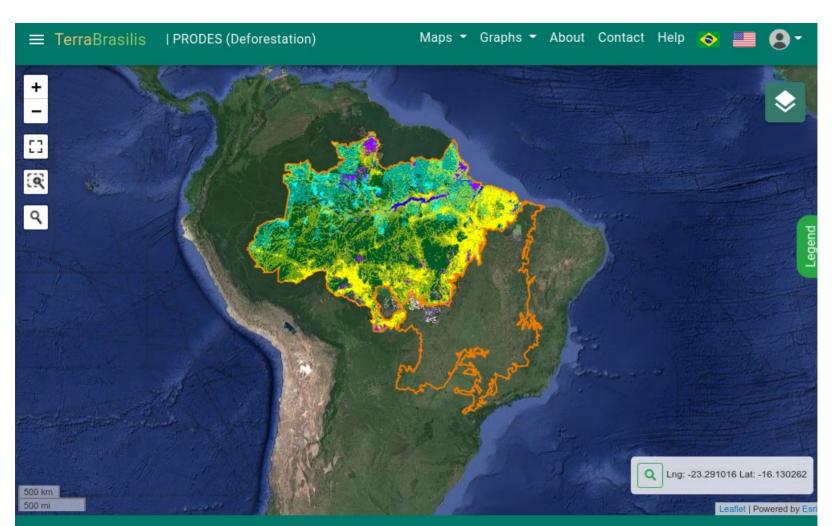
http://www.seer.ufu.br/index.php/braziliangeojournal/article/view/24886

Problema

O PRODES é um projeto desenvolvido pelo INPE e mapeia os polígonos de desmatamento por fotointerpretação das imagens de satélite por especialistas, e fornece as taxas de desmatamento uma vez por ano (INPE, 2019)







Solução

Para agilizar a quantificação do desmatamento na Amazônia, pode-se utilizar o conhecimento dos especialistas do PRODES para treinar algoritmos de Machine Learning.

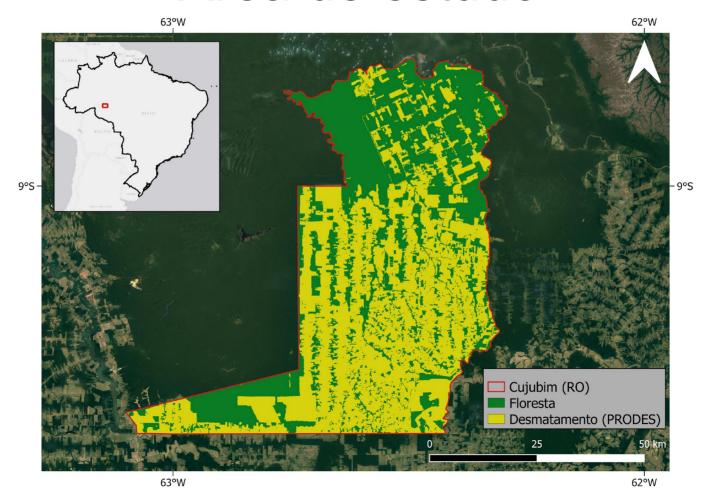


Objetivos

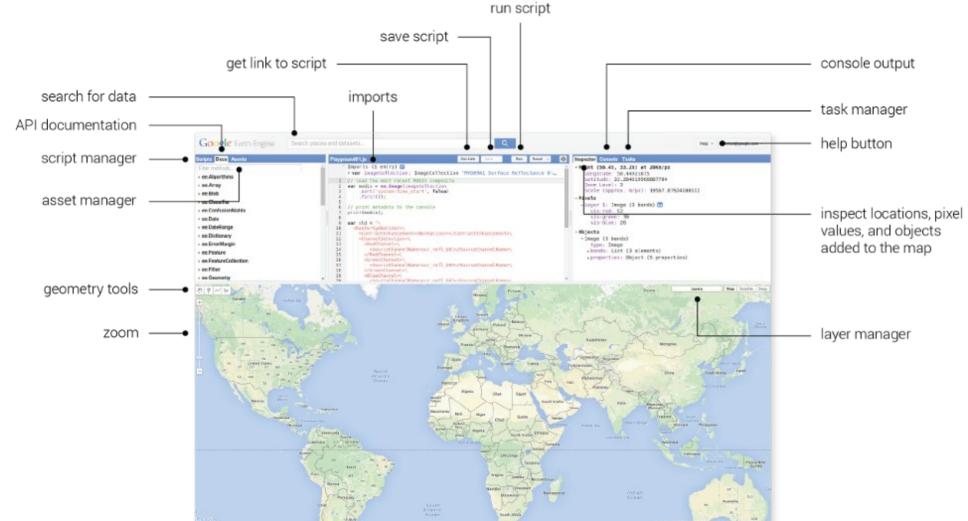
- 1) Classificar desmatamento em séries temporais (2015-2021) de imagens de satélite (Landsat 8/OLI) a partir de um algoritmo de Machine Learning treinado com amostras do PRODES.
- 2) Avaliar e comparar o desempenho de diferentes algoritmos.

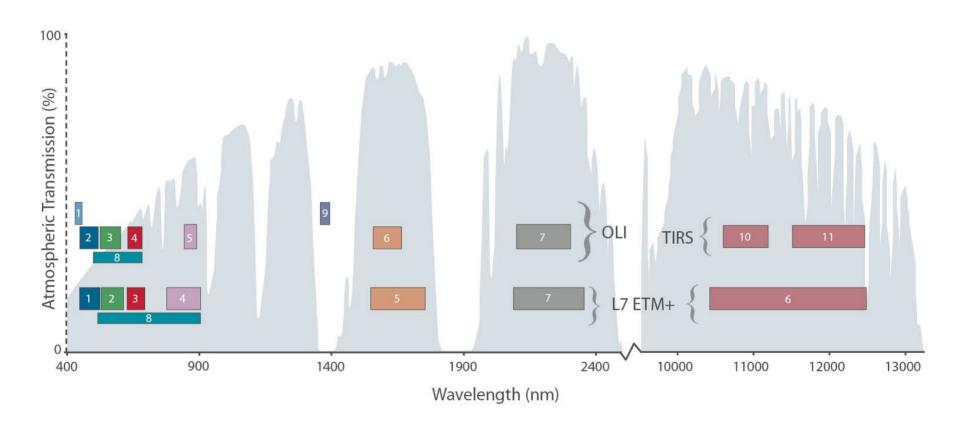
RandomForestClassifier, LogisticRegression, XGBClassifier, KNeighborsClassifier, BaggingClassifier, ExtraTreesClassifier, SGDClassifier, SVC, NuSVC, LinearSVC, BernoulliNB, LGBMClassifier, MLPClassifier, AdaBoostClassifier.

Área de estudo

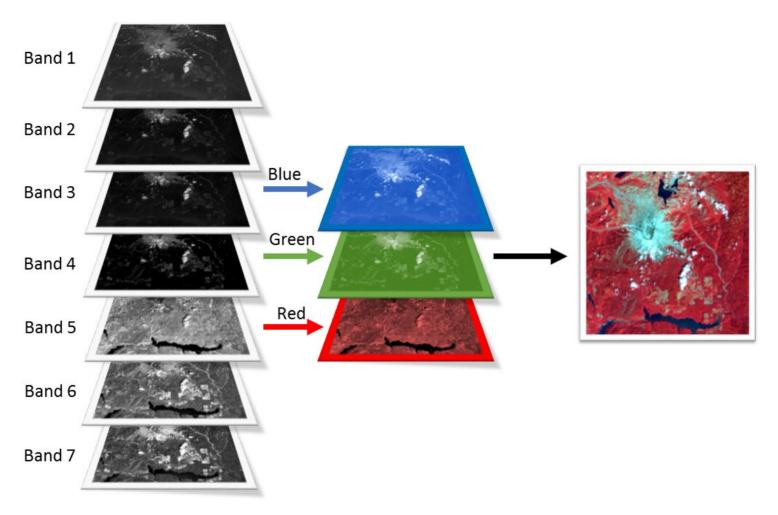


Processamento de Imagens (Google Earth Engine)



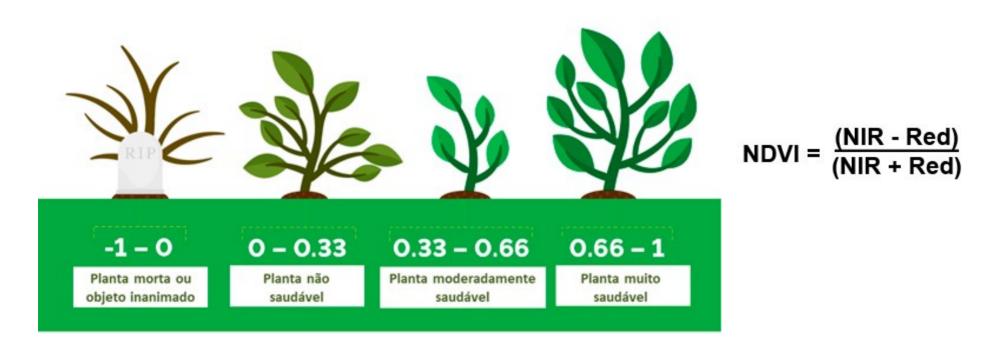


https://landsat.gsfc.nasa.gov/satellites/landsat-8/



https://gisenglish.geojamal.com/2019/10/band-combinations-for-landsat-8.html

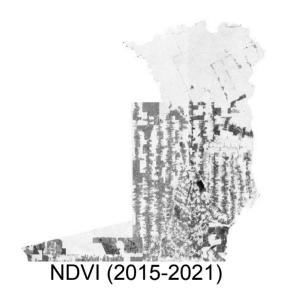
Normalized Difference Vegetation Index



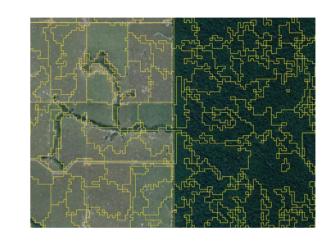
https://www.revistacanavieiros.com.br/o-que-sao-mapas-ndvi-e-como-utiliza-los-na-fazenda



Extração de atributos a partir dos valores dos pixels (em linguagem R)



Segmentação



Valores dos pixels extraídos por segmento (NDVI, NDWI e banda do SWIR)

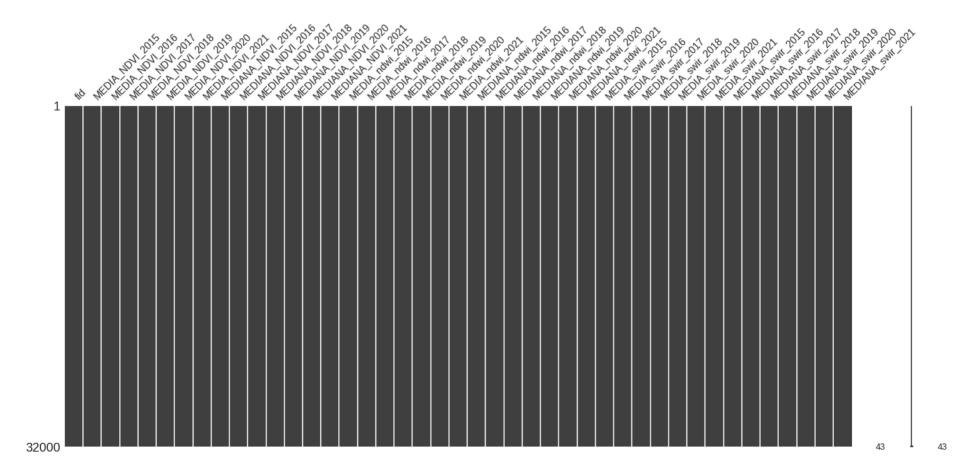
SHAPE: (106346, 85)

	masked	mean	classe_amostra	fid	MEDIA_NDVI_2015	MEDIA_NDVI_2016	MEDIA_NDVI_2017	MEDIA_NDVI_2018	MEDIA_NDVI_2019	MEDIA_NDVI_2020
0	-1078580800	-1078580800	0.0	1	0.860178	0.862347	0.880624	0.876968	0.874338	0.866653
1	-1127396665	-1127396665	NaN	2	0.826198	0.861789	0.858303	0.876121	0.843605	0.857290
2	231191332	231191332	0.0	3	0.835334	0.842979	0.841535	0.840028	0.818171	0.834221
3	-1127396665	-1127396665	NaN	4	0.840929	0.840058	0.832866	0.846502	0.839763	0.830259
4	231191332	231191332	0.0	5	0.836446	0.852750	0.864893	0.860996	0.865849	0.846589

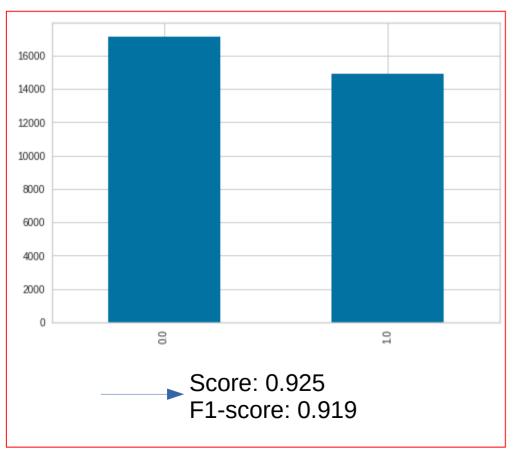
Cada linha da tabela corresponde a um segmento

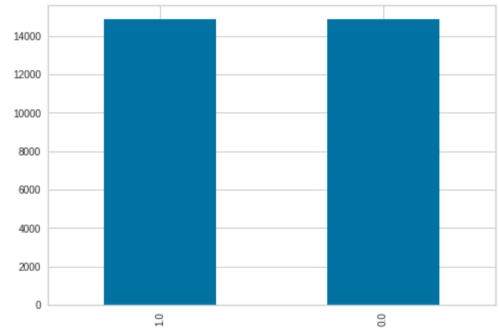
5 rows x 85 columns

Amostras de treinamento: linhas na tabela que possuem rótulo de classe 0 = Floresta e 1 = Desmatamento



Resultados (teste de balanceamento de classes com Random Forest) Train-test split de 30%





Score: 0.918

F1-score: 0.918

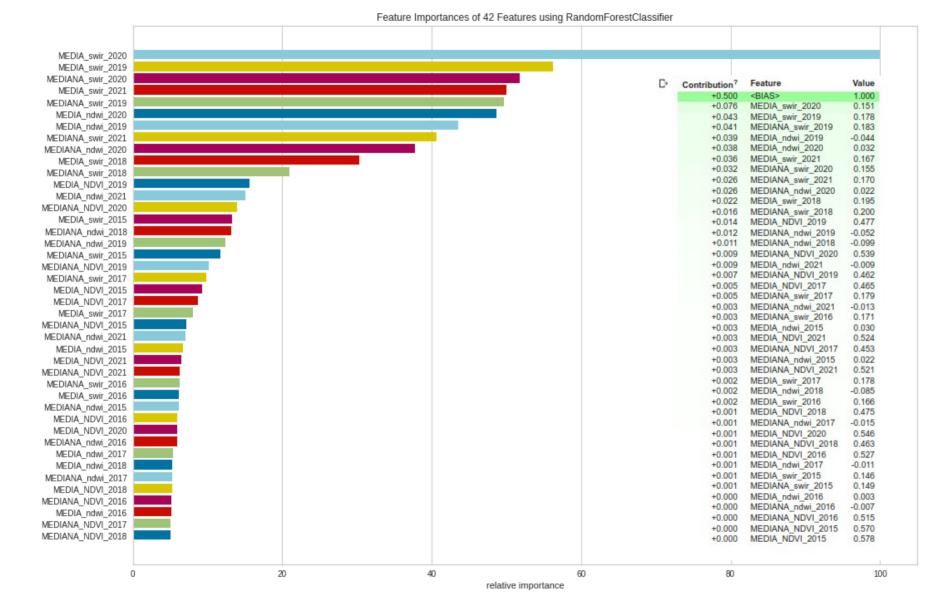
Resultados (testes de seleção de atributos com Random Forest)

0.25

0.00

-0.25

MEDIA_NDVI_2015	1	0.92	0.91	0.89	0.88	0.83	0.8	0.96	0.9	0.88	0.87	0.86	0.8	0.77	-0.95	-0.87	-0.86	-0.85	-0.84	-0.78	-0.76
MEDIA_NDVI_2016	0.92	1	0.91	0.9	0.88	0.83	0.81	0.87	0.95	0.86	0.85	0.83	0.78	0.76	-0.86	-0.92	-0.84	-0.83	-0.82	-0.76	-0.75
MEDIA_NDVI_2017	0.91	0.91	1	0.92	0.91	0.85	0.84	0.88	0.89	0.96	0.9	0.88	0.82	0.8	-0.86	-0.86	-0.93	-0.87	-0.85	-0.79	-0.78
MEDIA_NDVI_2018	0.89	0.9	0.92	1	0.93	0.87	0.85	0.86	0.86	0.88	0.96	0.88	0.83	0.8	-0.84	-0.84	-0.85	-0.93	-0.86	-0.8	-0.79
MEDIA_NDVI_2019	0.88	0.88	0.91	0.93	1	0.9	0.89	0.86	0.87	0.88	0.91	0.96	0.87	0.85	-0.85	-0.85	-0.86	-0.88	-0.94	-0.85	-0.84
MEDIA_NDVI_2020	0.83	0.83	0.85	0.87	0.9	1	0.89	0.79	0.8	0.81	0.83	0.85	0.94	0.83	-0.78	-0.78	-0.79	-0.81	-0.83	-0.91	-0.82
MEDIA_NDVI_2021	0.8	0.81	0.84	0.85	0.89	0.89	1	0.79	0.8	0.82	0.83	0.86	0.87	0.95	-0.77	-0.78	-0.8	-0.81	-0.84	-0.85	-0.93
MEDIA_ndwi_2015	0.96	0.87	0.88	0.86	0.86	0.79	0.79	1	0.91	0.91	0.89	0.89	0.82	0.8	-0.98	-0.9	-0.9	-0.88	-0.87	-0.81	-0.8
MEDIA_ndwi_2016	0.9	0.95	0.89	0.86	0.87	0.8	0.8	0.91	1	0.9	0.89	0.87	0.82	0.8	-0.89	-0.97	-0.89	-0.87	-0.87	-0.8	-0.8
MEDIA_ndwi_2017	0.88	0.86	0.96	0.88	0.88	0.81	0.82	0.91	0.9	1	0.92	0.91	0.84	0.83	-0.89	-0.88	-0.97	-0.9	-0.89	-0.83	-0.82
MEDIA_ndwi_2018	0.87	0.85	0.9	0.96	0.91	0.83	0.83	0.89	0.89	0.92	1	0.93	0.86	0.85	-0.88	-0.87	-0.9	-0.97	-0.91	-0.84	-0.83
MEDIA_ndwi_2019	0.86	0.83	0.88	0.88	0.96	0.85	0.86	0.89	0.87	0.91	0.93	1	0.89	0.88	-0.87	-0.86	-0.89	-0.9	-0.97	-0.88	-0.87
MEDIA_ndwi_2020	0.8	0.78	0.82	0.83	0.87	0.94	0.87	0.82	0.82	0.84	0.86	0.89	1	0.89	-0.81	-0.81	-0.83	-0.85	-0.88	-0.96	-0.88
MEDIA_ndwi_2021	0.77	0.76	0.8	0.8	0.85	0.83	0.95	0.8	0.8	0.83	0.85	0.88	0.89	1	-0.79	-0.79	-0.82	-0.83	-0.87	-0.87	-0.97
MEDIA_swir_2015	-0.95	-0.86	-0.86	-0.84	-0.85	-0.78	-0.77	-0.98	-0.89	-0.89	-0.88	-0.87	-0.81	-0.79	1	0.92	0.91	0.89	0.88	0.83	0.81
MEDIA_swir_2016	-0.87	-0.92	-0.86	-0.84	-0.85	-0.78	-0.78	-0.9	-0.97	-0.88	-0.87	-0.86	-0.81	-0.79	0.92	1	0.91	0.9	0.89	0.83	0.82
MEDIA_swir_2017	-0.86	-0.84	-0.93	-0.85	-0.86	-0.79	-0.8	-0.9	-0.89	-0.97	-0.9	-0.89	-0.83	-0.82	0.91	0.91	1	0.93	0.91	0.86	0.85
MEDIA_swir_2018	-0.85	-0.83	-0.87	-0.93	-0.88	-0.81	-0.81	-0.88	-0.87	-0.9	-0.97	-0.9	-0.85	-0.83	0.89	0.9	0.93	1	0.93	0.87	0.86
MEDIA_swir_2019	-0.84	-0.82	-0.85	-0.86	-0.94	-0.83	-0.84	-0.87	-0.87	-0.89	-0.91	-0.97	-0.88	-0.87	0.88	0.89	0.91	0.93	1	0.91	0.89
MEDIA_swir_2020	-0.78	-0.76	-0.79	-0.8	-0.85	-0.91	-0.85	-0.81	-0.8	-0.83	-0.84	-0.88	-0.96	-0.87	0.83	0.83	0.86	0.87	0.91	1	0.91
MEDIA_swir_2021	-0.76	-0.75	-0.78	-0.79	-0.84	-0.82	-0.93	-0.8	-0.8	-0.82	-0.83	-0.87	-0.88	-0.97	0.81	0.82	0.85	0.86	0.89	0.91	1
	2015	16	2017	2018	2019	2020	2021	2015	2016	2017	2018	2019	2020	121	2015	2016	2017	2018	2019	2020	2021
		NDVI_2016				1_20						vi_20	и_20	ndwi_2021							
	NDM	N N	NDM	NDM	NDM	NDM	NDV.	ndwi	ndwi	ndwi	ndwi	ndwi	ndwi		SWif	SWir	swir	SWIF	SWir	SWir	SWir
		WEDIA	MEDIA	MEDIA	ÆDIA	ÆDIA	ÆDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA	MEDIA
	MEDIA	Æ	¥	W	Æ	Æ	Æ	띨	띨	핃	띨	띨	\exists	띨	3	N	3	3	8	¥	N



Resultados (testes de seleção de atributos com Random Forest)

Train-test split de 30%

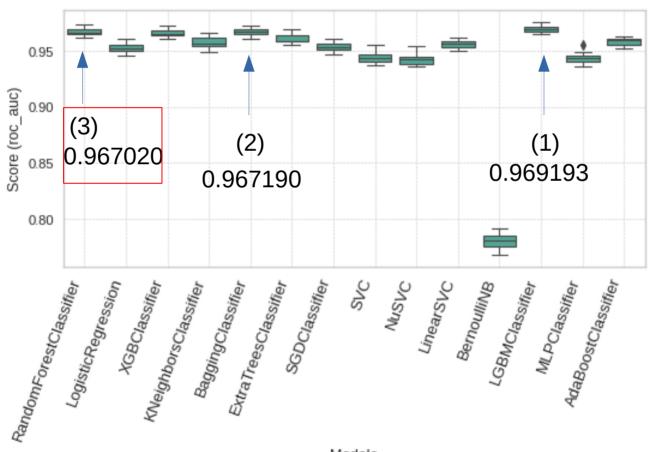
Entrada	Score	F1-Score
Todos os atributos (média e mediana de NDVI, NDWI e SWIR)	0.9245	0.919
Atributos de média e mediana de NDVI e NDWI	0.916	0.909
Atributos de média e mediana de SWIR	0.921	0.915
Atributos com contribuição > 0.001	0.9246	0.919

Decidimos utilizar todos os atributos 42 no total

Resultados (Escolha do modelo)

Busca de Hiperparâmetros com HalvingGridSearchCV com métrica roc_auc

RepeatedStratifiedKFold: 10 folds

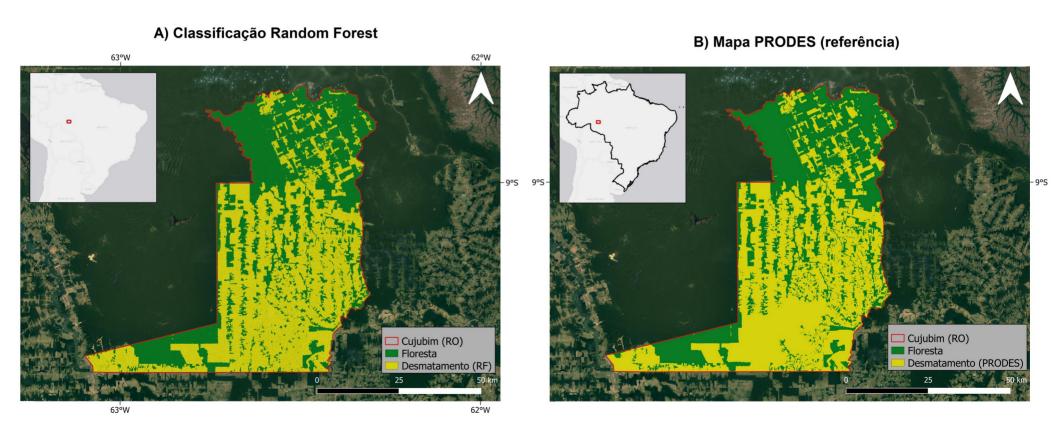


Resultados (BayesSearchCV e Random Forest)



Melhor estimador	Best score (roc auc)
max_depth = 26 max_samples = 1.0 min_samples_leaf = 3 n_estimators= 595	~0.97

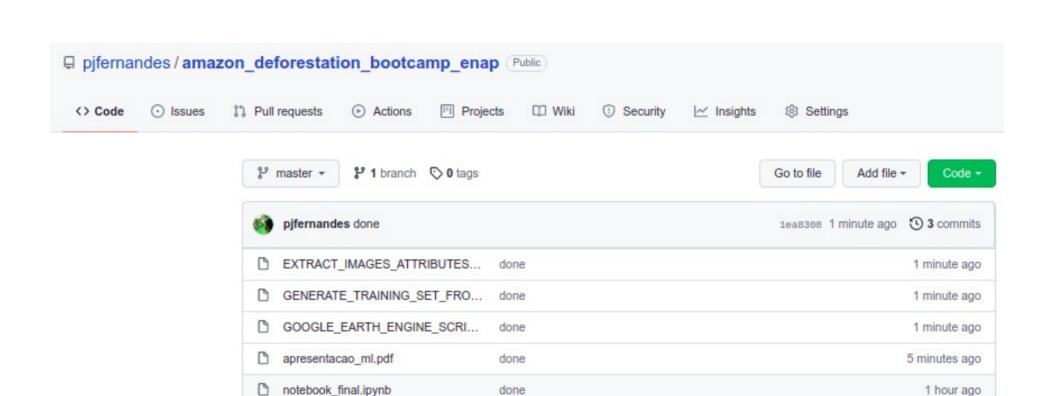
Modelo em "produção"



Considerações finais

1) Modelo escolhido foi o Random Forest pelo bom desempenho e pelo uso com bons resultados na literatura de Sensoriamento Remoto.

- 2) O objetivo não é substituir o PRODES, e sim fornecer valores de área de desmatamento enquanto os valores anuais do PRODES não são divulgados.
- 3) Além do notebook com os testes de machine learning, outras contribuições foram:
 - O desenvolvimento de códigos para processamento de imagens e segmentação no Google Earth Engine
 - Extração de atributos das imagens na linguagem R, contribuindo com o uso de OBIA (Object Based Image Analysis) sem depender de softwares pagos



Help people interested in this repository understand your project by adding a README.

Add a README

Planos futuros

1) Testar o modelo em áreas em que não houve coleta de amostras.

2) Publicar artigo científico.

3) Aprofundamento em Deep Learning e em visão computacional.

Identificação de desmatamento na Amazônia em imagens de satélite

Cassiano Zago (DEPEN)
Pedro J. Fernandes (UFF)
Washington R. Dias (Prefeitura de Vitória – ES)

https://github.com/pjfernandes/amazon_deforestation_bootcamp_enap