

# **Identificação de desmatamento na Amazônia em imagens de satélite**

Cassiano Zago (DEPEN)

Pedro J. Fernandes (UFF)

Washington R. Dias (Prefeitura de Vitória – ES)

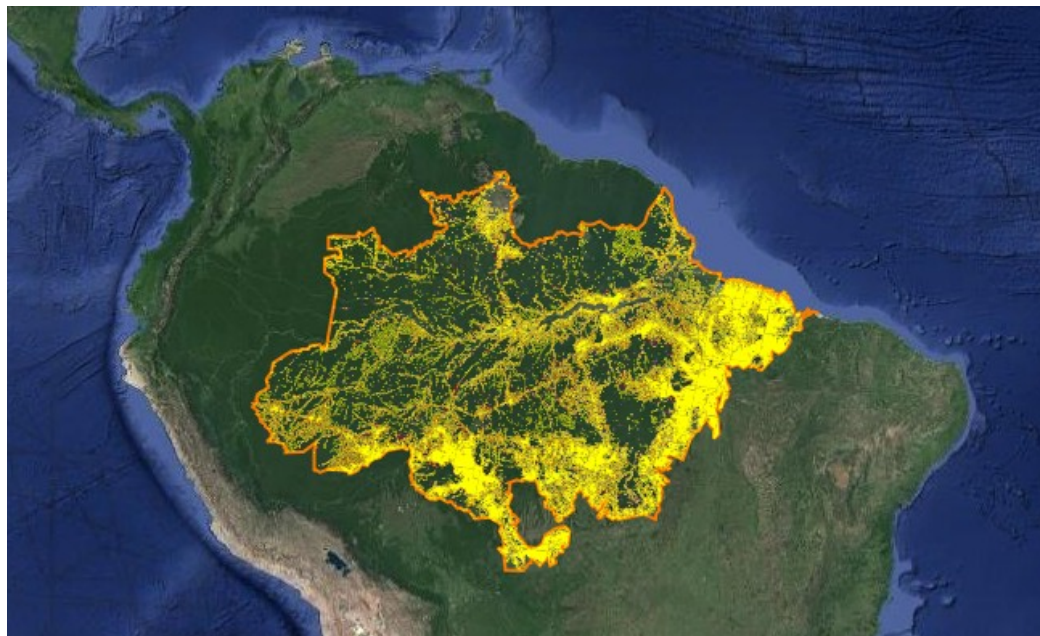
[https://github.com/pjfernandes/amazon\\_deforestation\\_bootcamp\\_enap](https://github.com/pjfernandes/amazon_deforestation_bootcamp_enap)

# 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

<http://www.guladigital.info/>

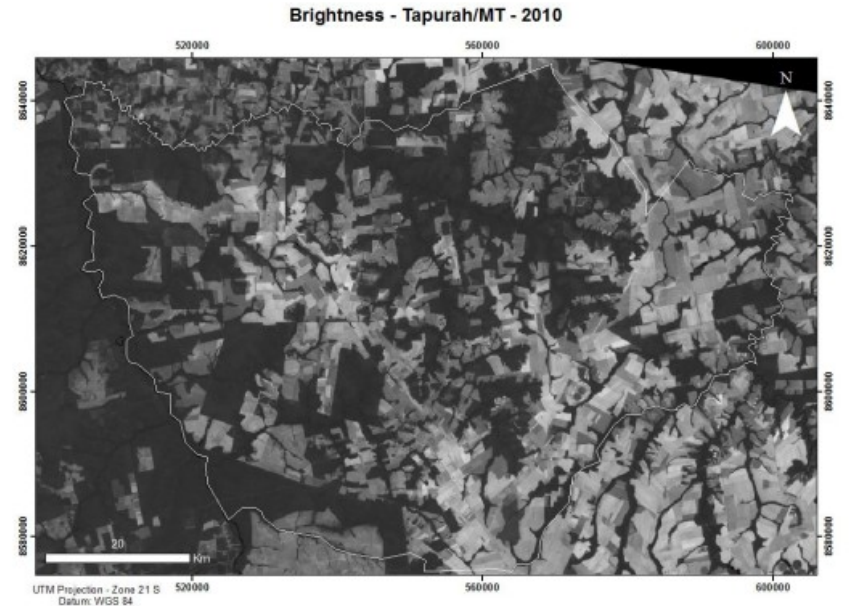


Alterações no clima local e regional

<https://redesustentabilidade.org.br>

# Introdução (contexto)

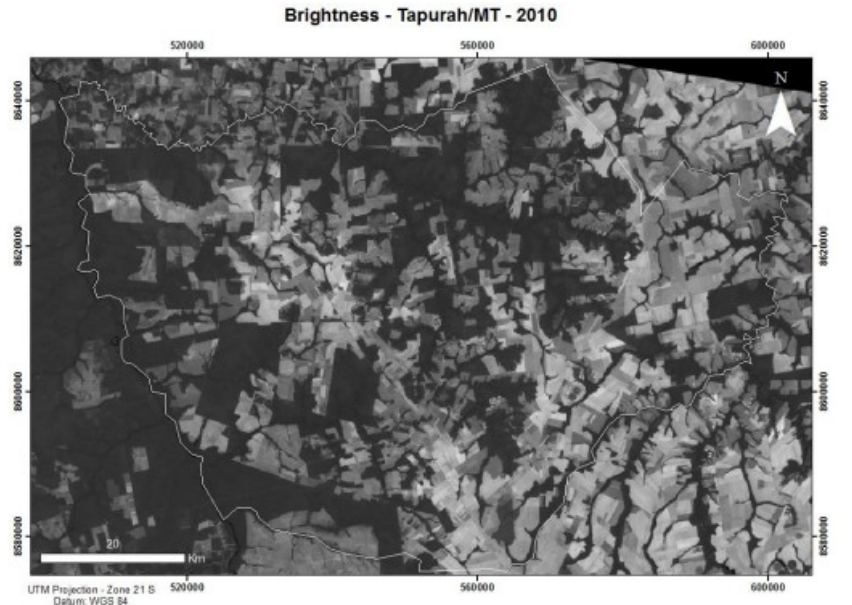
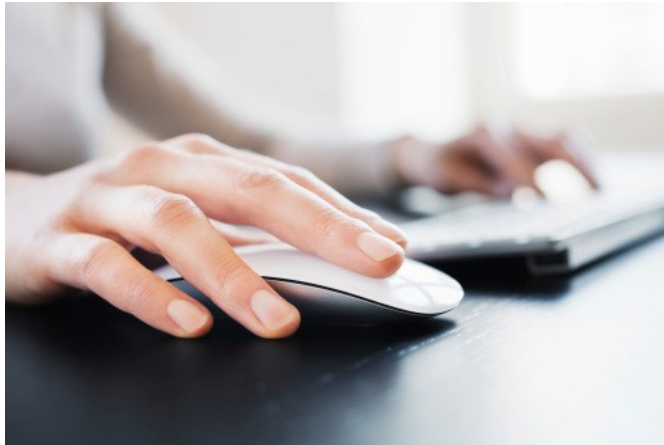
Monitoramento do desmatamento por imagens de Sensoriamento Remoto



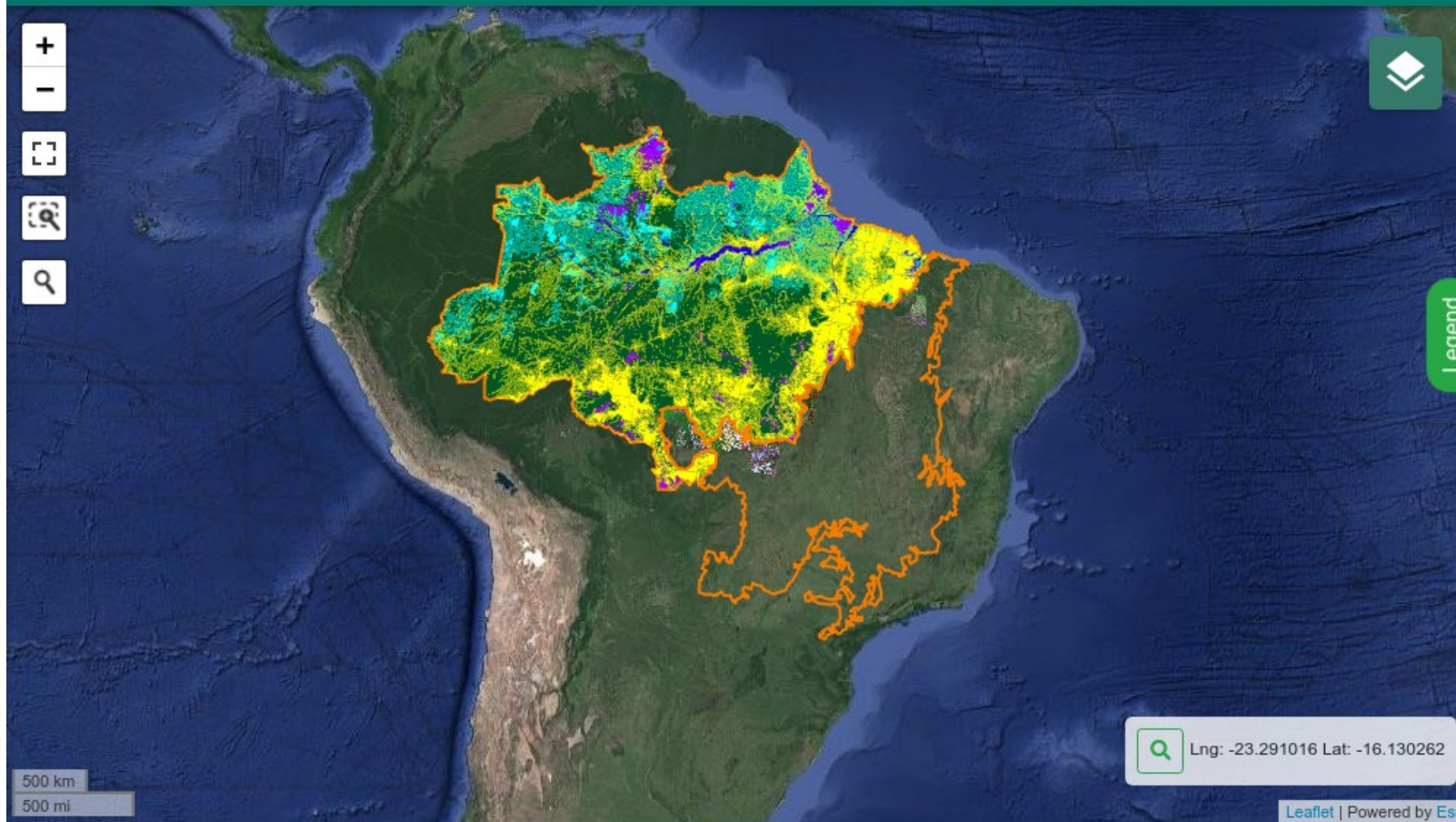
<http://www.seer.ufu.br/index.php/braziliangeojournal/article/view/24886>

# 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

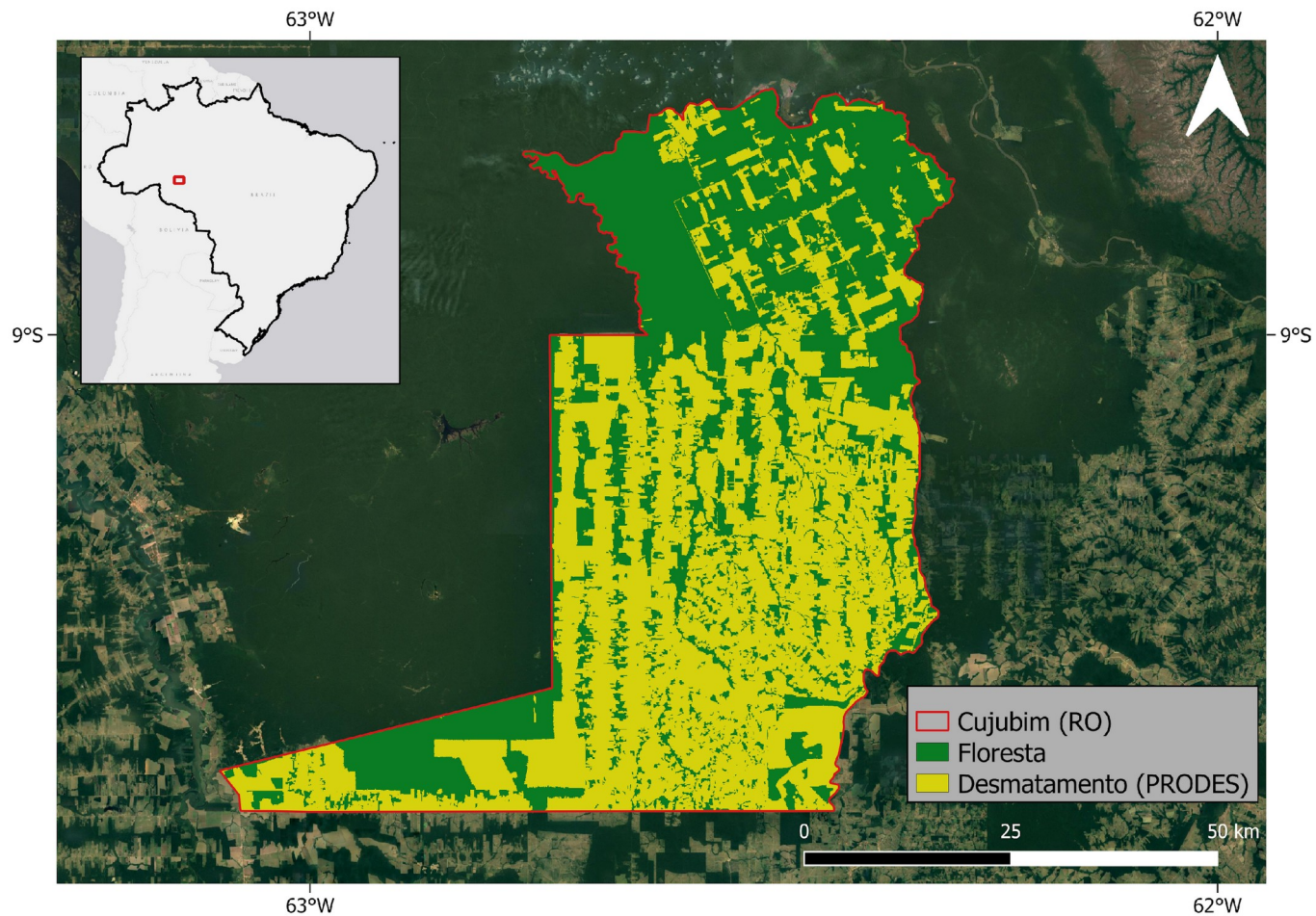
1) Classificar desmatamento em séries temporais (2015-2021) de imagens de satélite (Landsat 8/OLI) a partir de um algoritmo de Machine Learning treinado com amostras do PRODES.

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

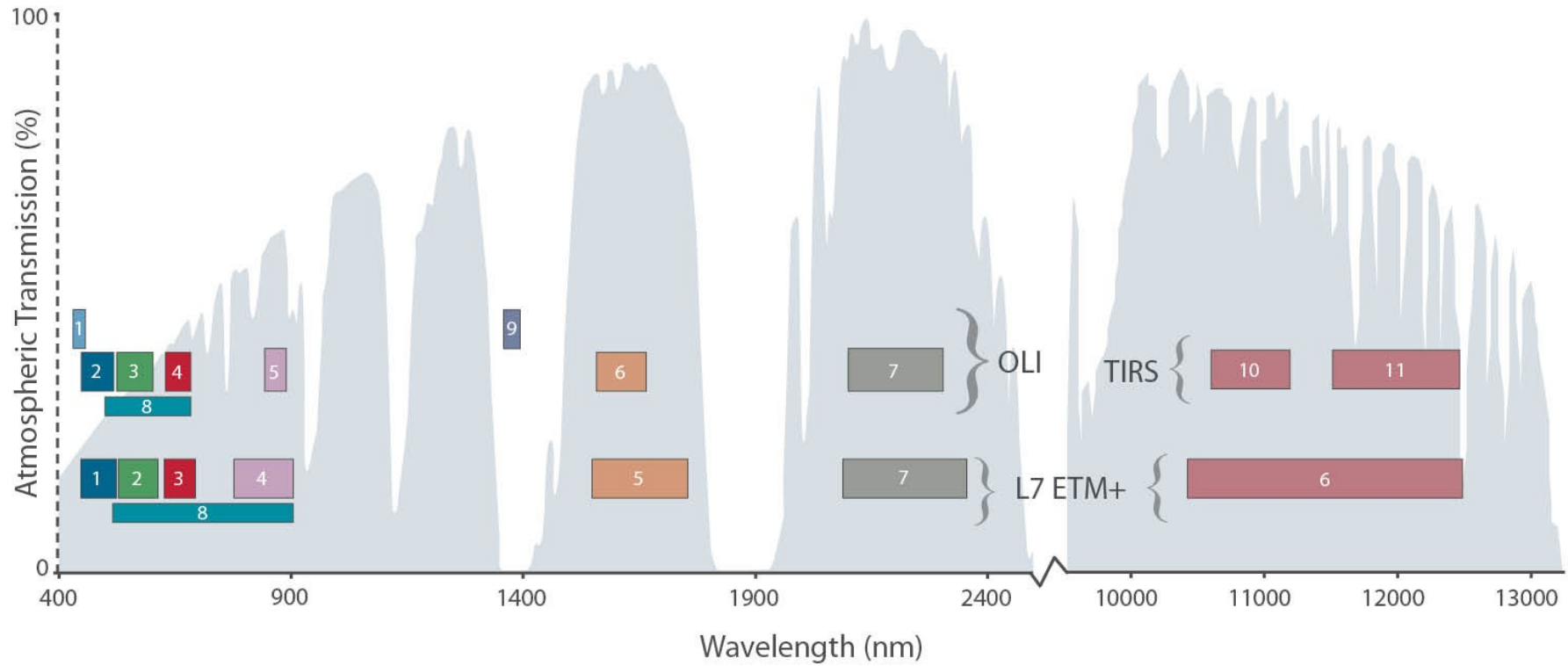


# Processamento de Imagens (Google Earth Engine)

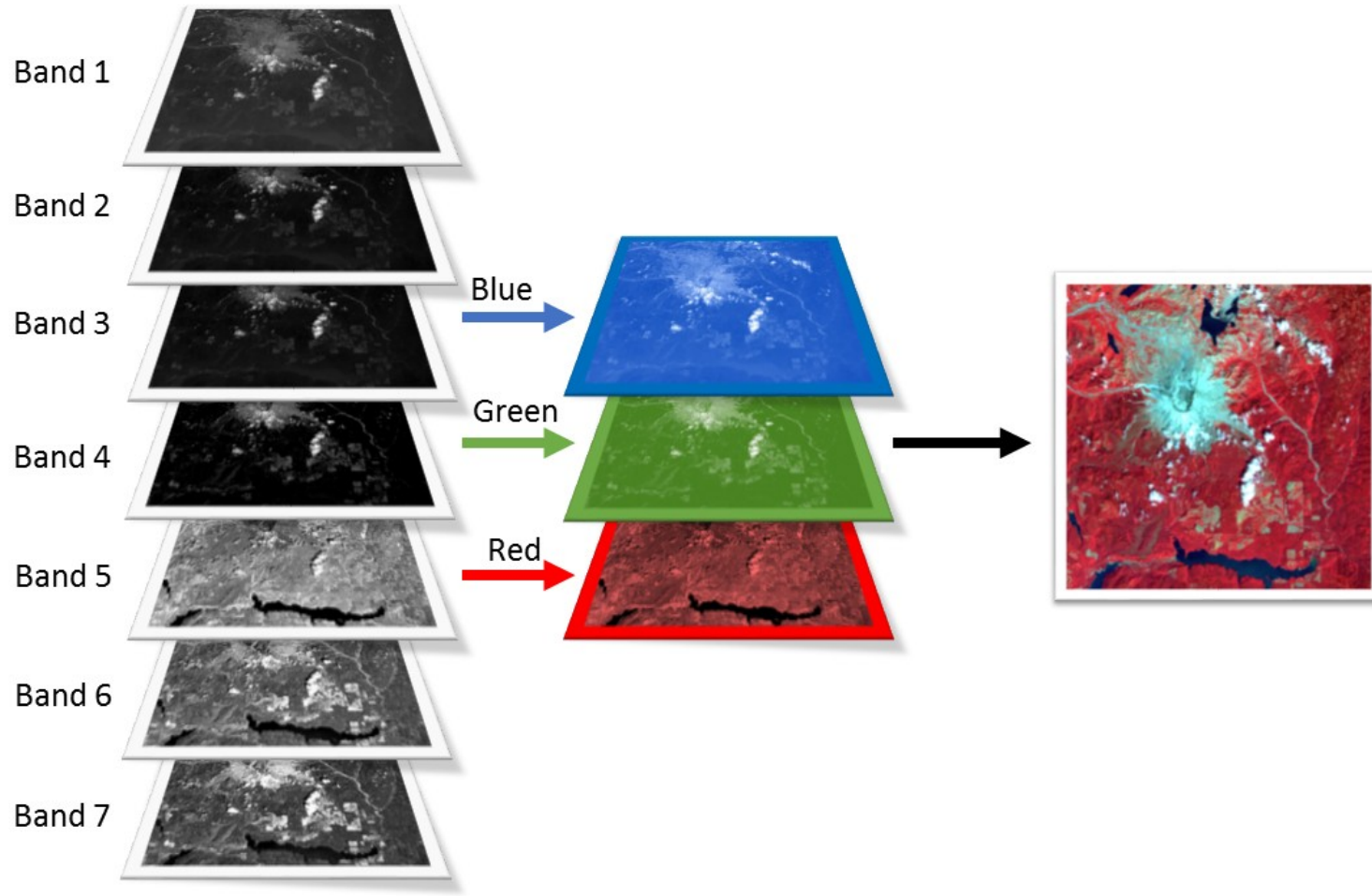
Diagram illustrating the Google Earth Engine interface and its components:

- search for data
- API documentation
- script manager
- asset manager
- geometry tools
- zoom
- run script
- save script
- get link to script
- imports
- console output
- task manager
- help button
- inspect locations, pixel values, and objects added to the map
- layer manager

The screenshot shows the Google Earth Engine web application. The top navigation bar includes 'Script', 'Data', and 'Assets' tabs. The 'Script' tab is active, displaying a code editor with a JavaScript script for processing satellite imagery. The script includes comments and code for loading data, setting a region of interest, and displaying the result. The 'Data' tab shows a search bar and a list of datasets. The 'Assets' tab shows a list of user assets. The bottom panel displays a map of the world with a region of interest highlighted in the Amazon basin. The right sidebar contains the 'Inspector' panel, which shows the details of the selected region, including its coordinates, pixel values, and objects. The 'Console' panel shows the output of the script execution. The 'Task Manager' panel shows the status of tasks. The 'Help' button is located in the top right corner. The 'Layer Manager' panel shows the layers added to the map.



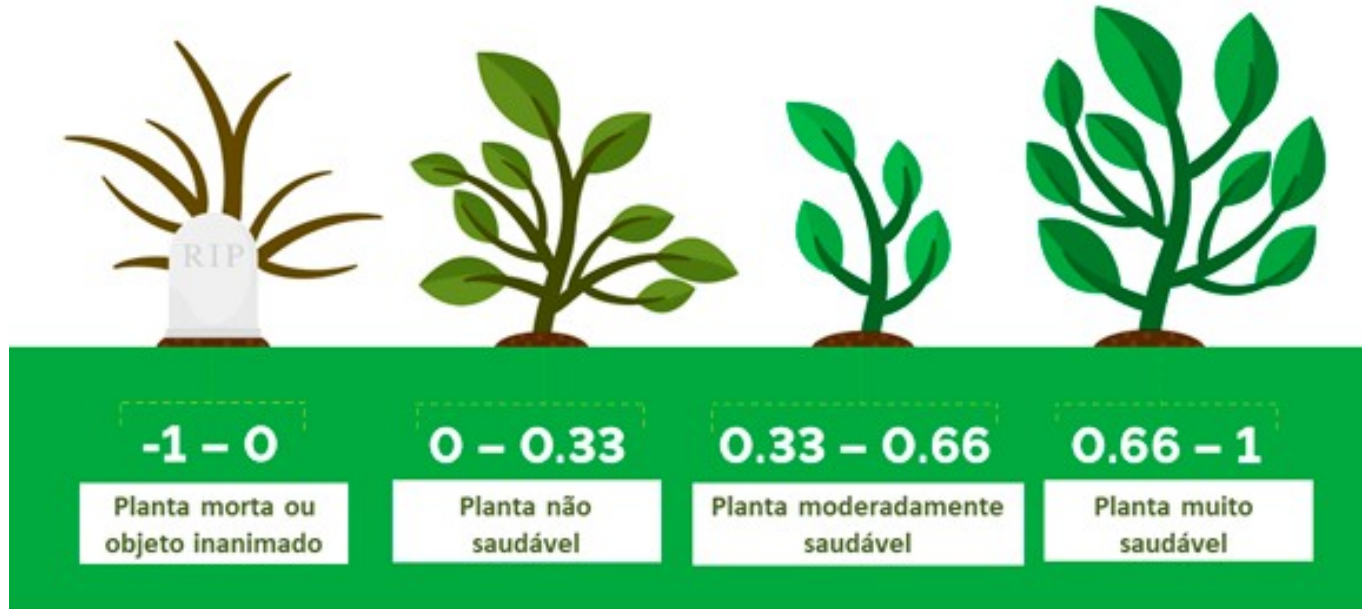
<https://landsat.gsfc.nasa.gov/satellites/landsat-8/>



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



$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})}$$



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





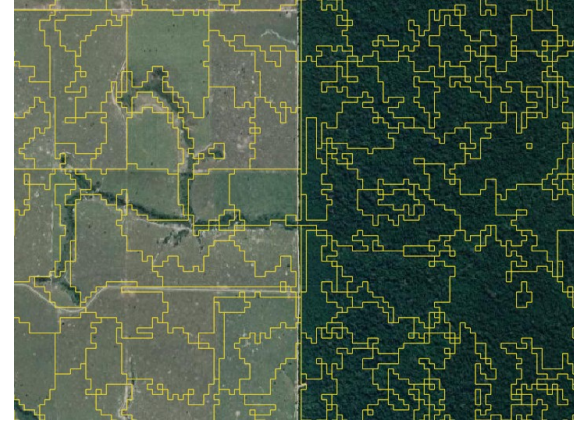
NDVI (Normalized  
Difference Vegetation  
Index) MODIS  
24/05/2016

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



NDVI (2015-2021)

Segmentação



Valores dos pixels extraídos por segmento

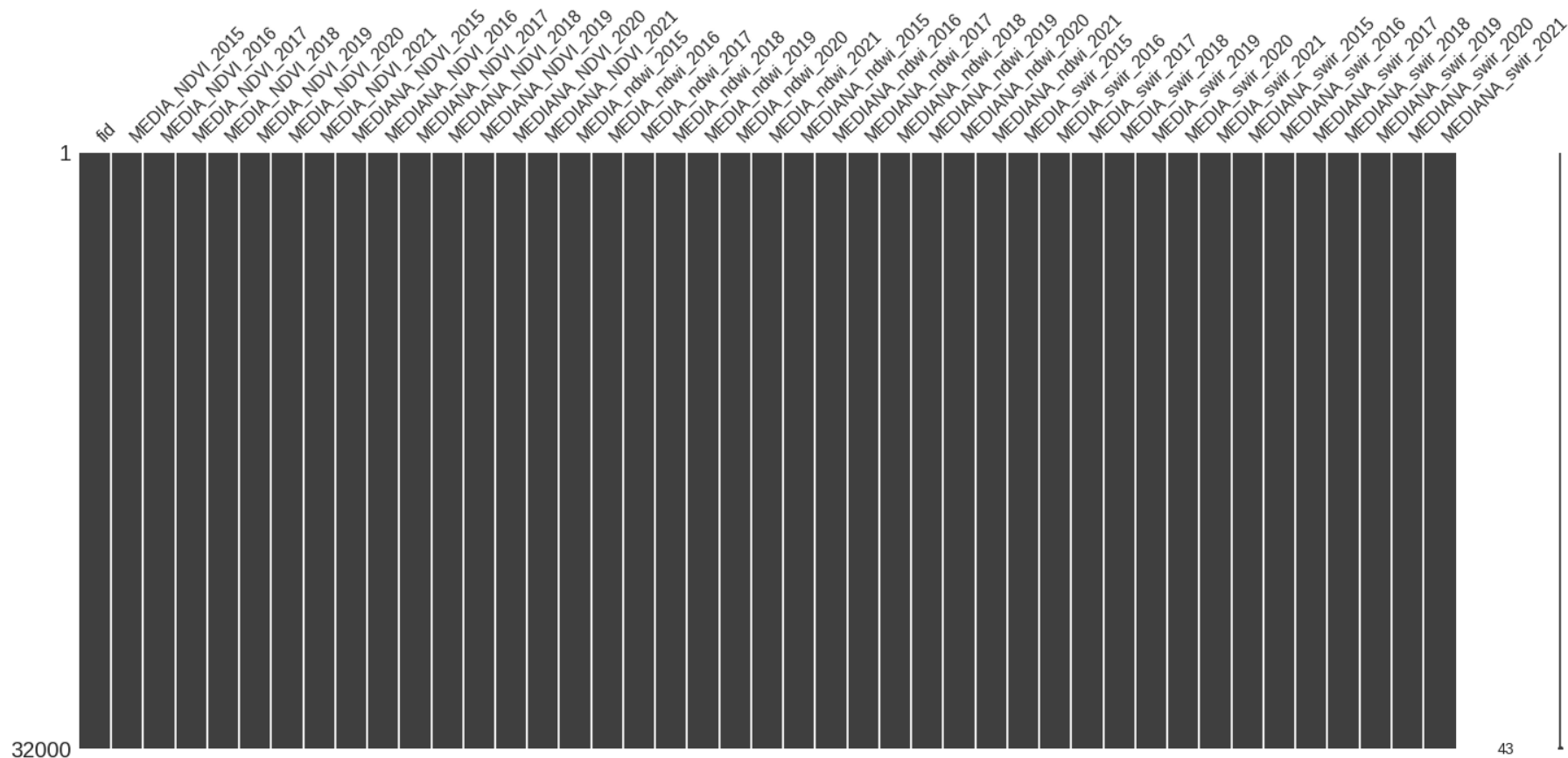
SHAPE: (106346, 85)

	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

5 rows x 85 columns

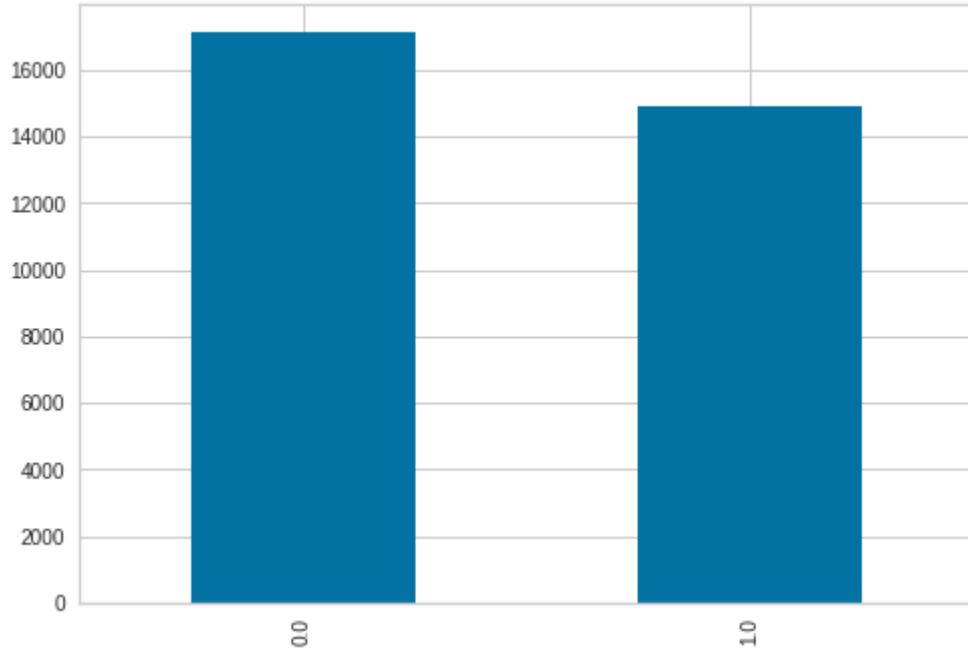
Cada linha da tabela corresponde a um segmento

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

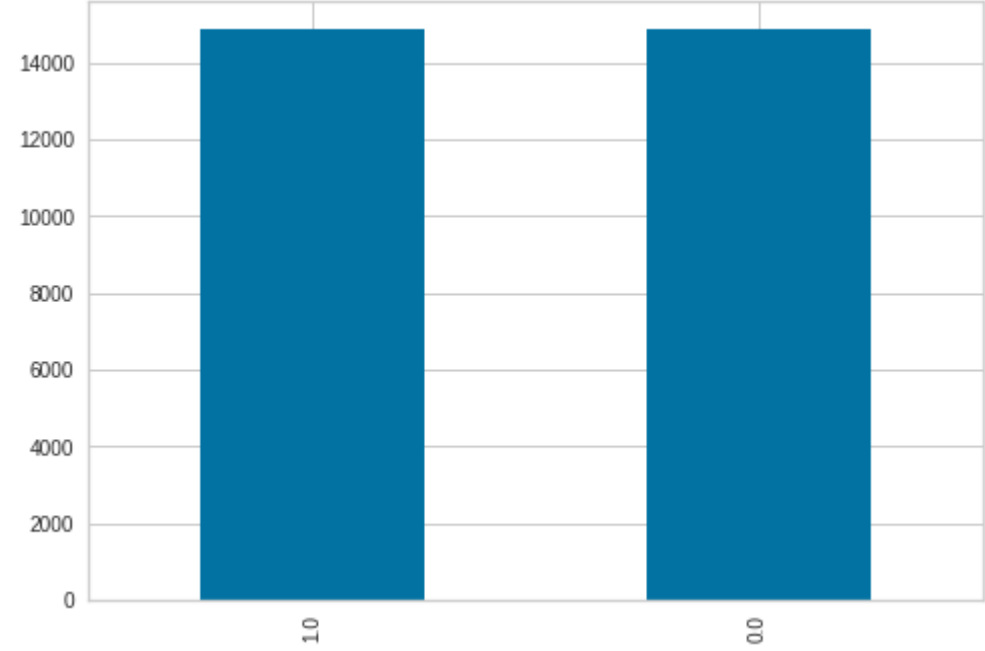


# Resultados (teste de balanceamento de classes com Random Forest)

Train-test split de 30%



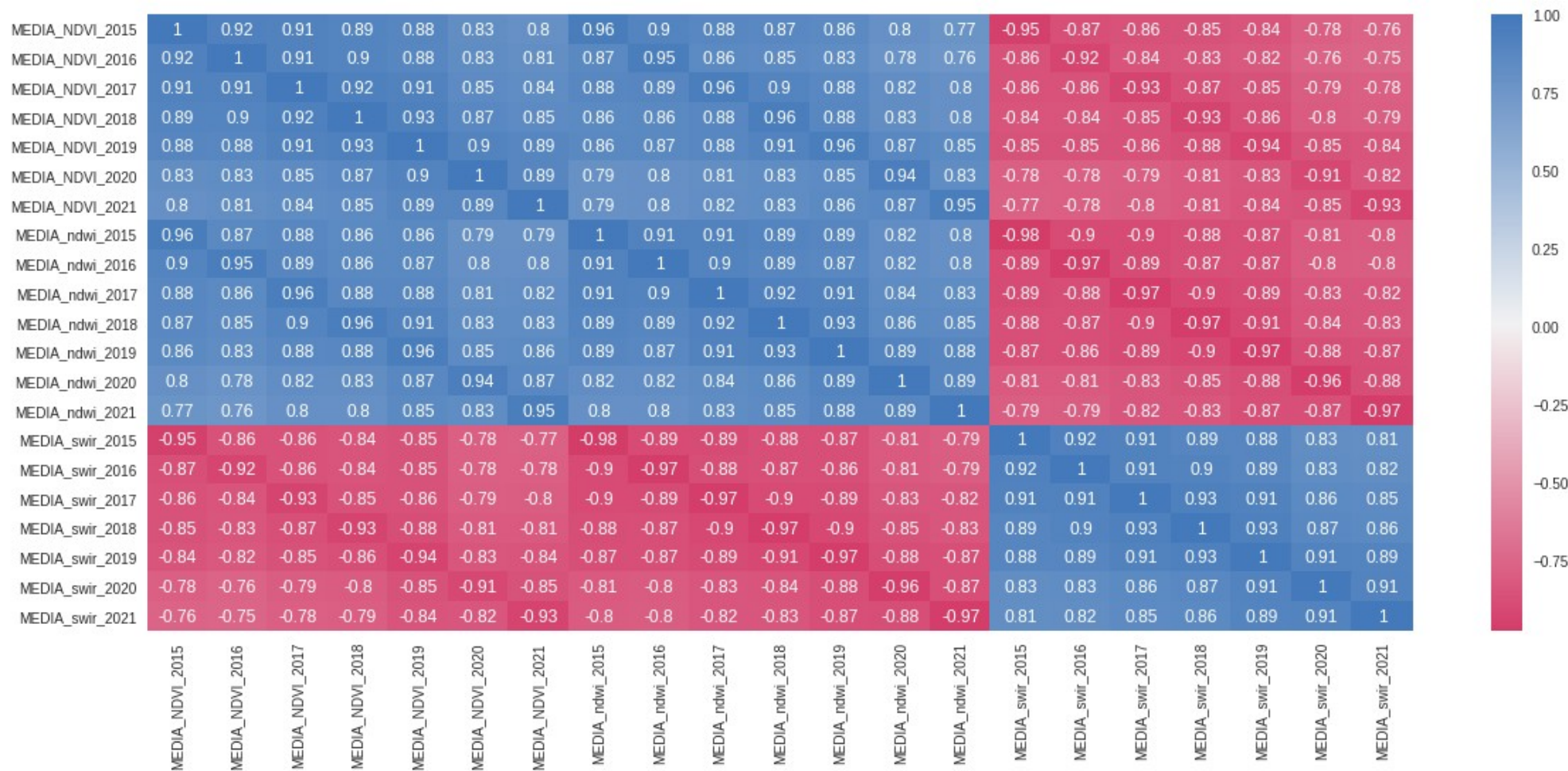
Score: 0.925  
F1-score: 0.919



Score: 0.918  
F1-score: 0.918

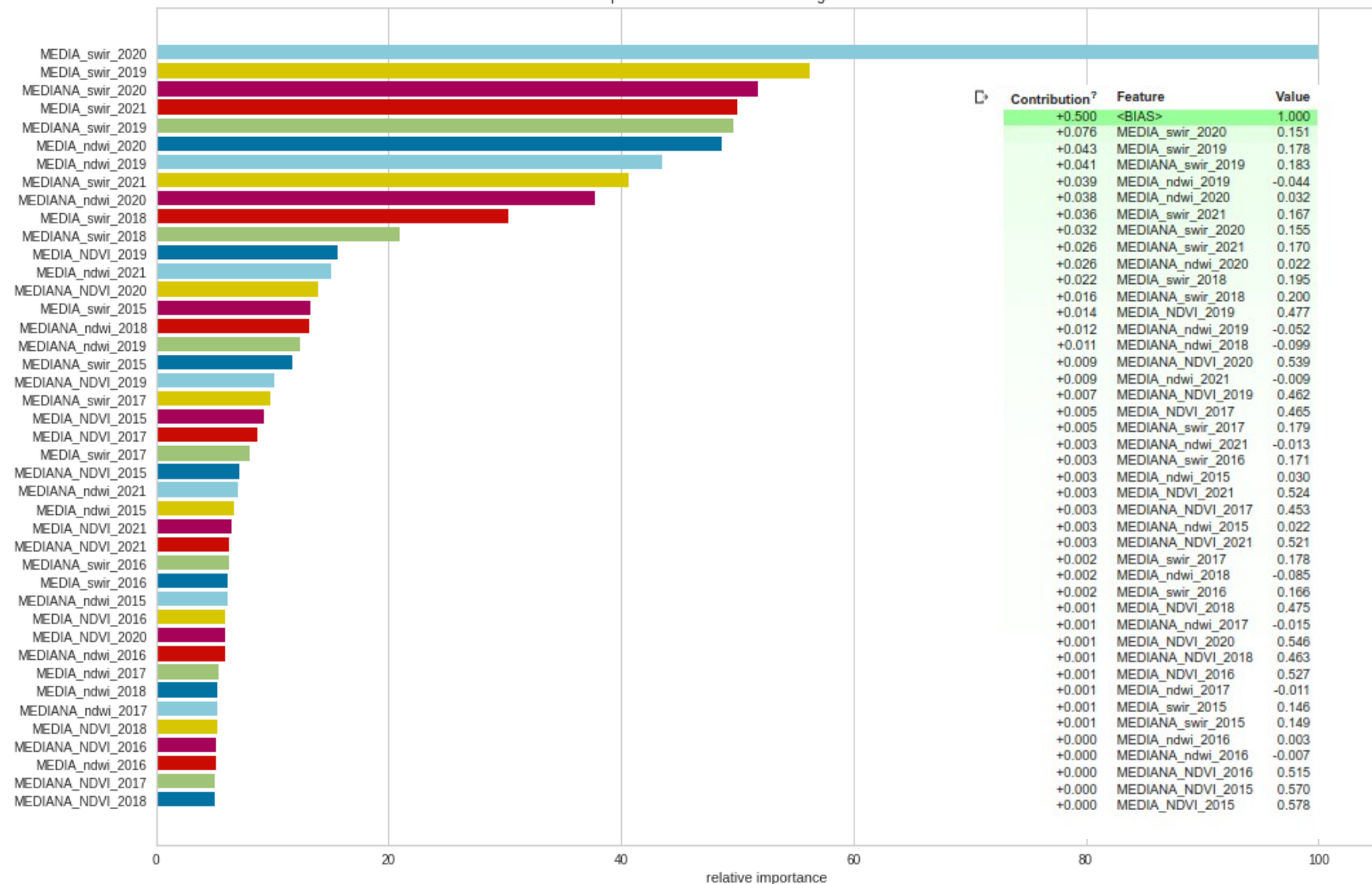


# Resultados (testes de seleção de atributos com Random Forest)





Feature Importances of 42 Features using RandomForestClassifier



# Resultados (testes de seleção de atributos com Random Forest)

Train-test split de 30%

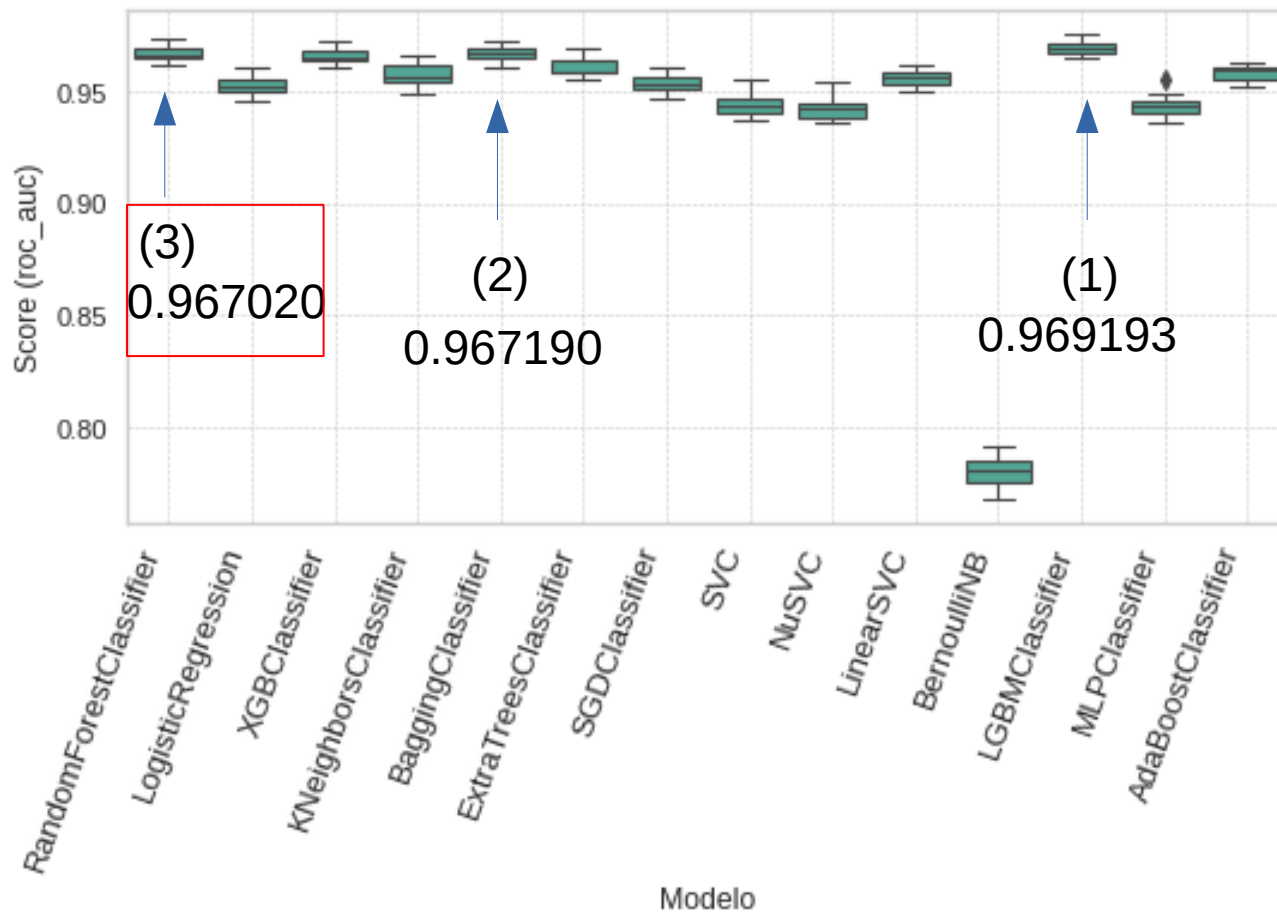
Entrada	Score	F1-Score
Todos os atributos (média e mediana de NDVI, NDWI e SWIR)	0.9245	0.919
Atributos de média e mediana de NDVI e NDWI	0.916	0.909
Atributos de média e mediana de SWIR	0.921	0.915
Atributos com contribuição > 0.001	0.9246	0.919

Decidimos utilizar todos os atributos  
42 no total

# Resultados (Escolha do modelo)

Busca de Hiperparâmetros com HalvingGridSearchCV com métrica roc\_auc

RepeatedStratifiedKFold: 10 folds



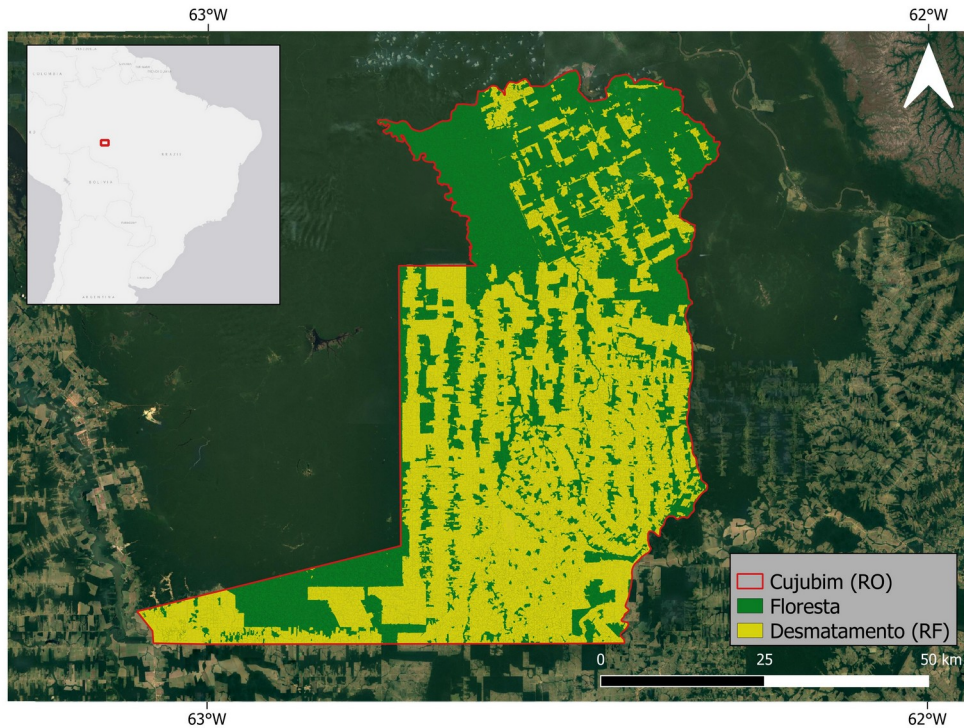
# Resultados (BayesSearchCV e Random Forest)



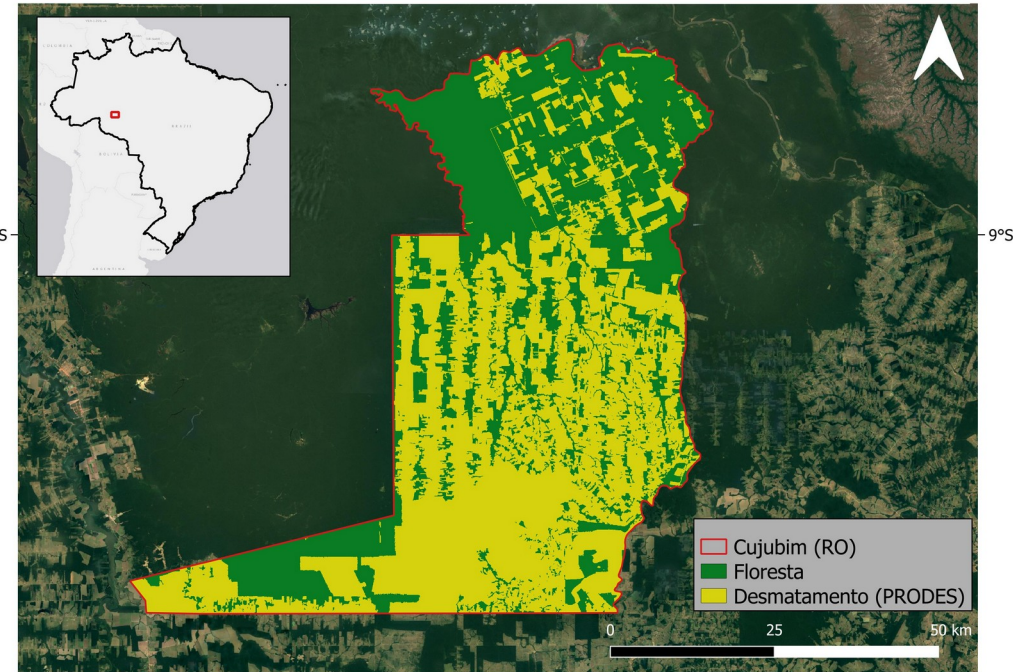
Melhor estimador	Best score (roc auc)
max_depth = 26 max_samples = 1.0 min_samples_leaf = 3 n_estimators= 595	~0.97

# Modelo em produção

A) Classificação Random Forest



B) Mapa PRODES (referência)





# 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

 **pjfernandes / amazon\_deforestation\_bootcamp\_enap** Public

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

 master  1 branch  0 tags


[Go to file](#)






[Add file](#)

[Code](#)



**pjfernandes** done

1ea8388 1 minute ago  3 commits

	EXTRACT_IMAGES_ATTRIBUTES...	done	1 minute ago
	GENERATE_TRAINING_SET_FRO...	done	1 minute ago
	GOOGLE_EARTH_ENGINE_SCRI...	done	1 minute ago
	apresentacao_ml.pdf	done	5 minutes ago
	notebook_final.ipynb	done	1 hour ago

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](https://github.com/pjfernandes/amazon_deforestation_bootcamp_enap)