

# Análise de metodologias para classificação de séries temporais aplicadas a agricultura

Leonardo Vieira  
Gerson Barbosa

Computação Aplicada - INPE

December 19, 2019

- 1 Introdução
- 2 Referencial Teórico
- 3 Sentinel-2 for Agriculture Monitoring (Sen2-Agri)
- 4 Motivação
- 5 Dados
- 6 Materiais e Métodos
- 7 Resultados

- As imagens de satélite tornaram-se uma importante fonte de dados que são utilizadas para realizar o monitoramento da terra;
- Os sensores a bordo dos satélites facilitam a varredura de grandes áreas inacessíveis em um curto período de tempo;
- Existem inúmeras aplicações nesse âmbito, sendo elas de grande importância na área ambiental, social e política;

- Uma importante informação que pode ser extraída das imagens de satélite são as séries temporais dos índices de vegetação tais como NDVI *Normalized Difference Vegetation Index*:

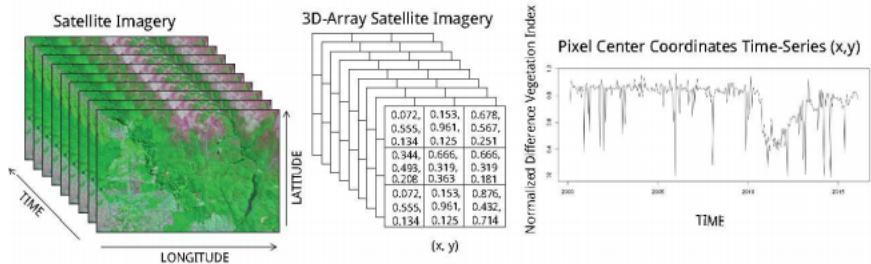
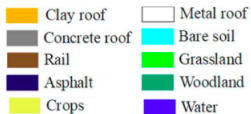


Figure: Série Remporal NDVI.

As séries temporais são utilizadas em aplicações que geram maps do uso e cobertura da terra.

# Uso e cobertura do solo

Land Cover: what is there?



Reference Image



Land Use: what is going on there?



Figure: Land cover and land use

# Random Forest (RF)

É definida como um classificador e regressor que consiste em uma coleção estruturada de árvores de decisão. A RF utiliza a técnica *bagging*, descrita como:

- Cria múltiplos conjuntos de dados selecionando aleatoriamente exemplos do conjunto original e de mesmo tamanho;
- Gera árvores de decisão para cada *Bootstrapped dataset* utilizando um subconjunto aleatório de características (*features*);
- Combina classificadores gerados (árvores de decisão) utilizando suas previsões (média, mediana ou moda);
- Utiliza os dados não selecionados (*out-of-bag dataset*) no passo 1 para avaliar a acurácia do classificador.

Uma visão geral dessa técnica é apresentada na figura a seguir:

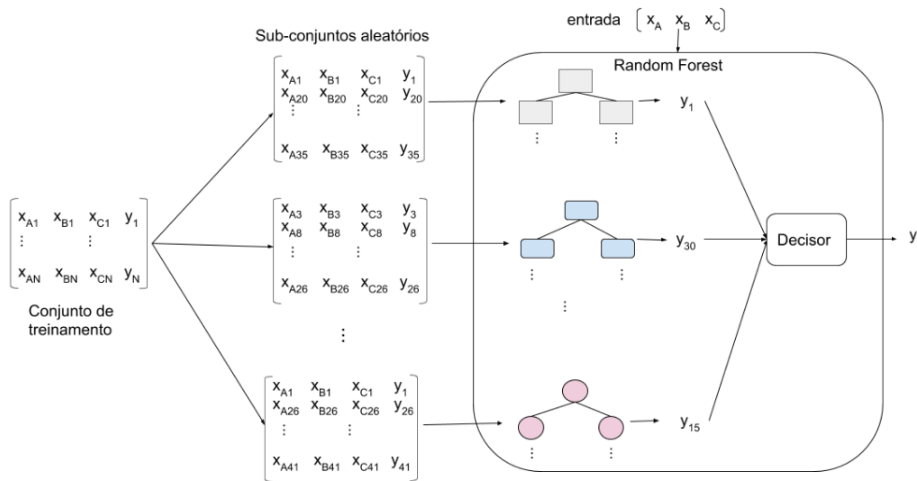


Figure: Visão Geral - *Random Forest*

São dados pelas combinações da reflectância espectral de dois ou mais comprimentos de onda para destacar uma característica relativa de interesse, como vegetação, características artificiais, hídricas e geológicas.

A fórmula do Índice de Vegetação por Diferença Normalizada (NDVI) é dada por:

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

A fórmula do Índice de Diferença Normalizada da Água é:

$$NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)}$$

A fórmula do brightness é dada por:

$$B = \sqrt{G^2 + R^2 + NIR^2 + SWIR^2}$$



# SITS: Data Analysis and Machine Learning for Data Cubes using Satellite Image Time Series

SITS (Satellite Image Time Series) é um pacote em R desenvolvido para realizar análises de séries temporais de sensoriamento remoto.

- Provê suporte a diversos métodos de aprendizado de máquinas, que incluem:
  - Análise de discriminação linear e quadrática;
  - SVM (Support Vector Machines);
  - Random Forest;
  - Neural Network.
- Aborda toda metodologia de classificação de séries temporais e sensoriamento remoto
  - aquisição de dados;
  - visualização;
  - clusterização;
  - etc;

sentinel-2  
sentinel-2 for agriculture monitoring

esa

sites products system overview dashboard custom jobs monitoring users data sources logout

Filter Reset Filter

- Bahia
  - L2A Atmospheric correction
  - L3A Composite product
  - L3B LAI mono-date product
  - L4A Crop mask product
    - S2AGRI\_L4A\_PRD\_S2\_20191106191917\_V20180902
    - S2AGRI\_L4A\_PRD\_S2\_201911061214754\_V20180902
    - S2AGRI\_L4A\_PRD\_S2\_201911061005353\_V20180902
    - S2AGRI\_L4A\_PRD\_S2\_201911061044700\_V20180902
    - S2AGRI\_L4A\_PRD\_S2\_201911061090513\_V20180902
    - S2AGRI\_L4A\_PRD\_S2\_201911061134529\_V20180902
    - S2AGRI\_L4A\_PRD\_S2\_201911061175008\_V20180902
  - L4B Crop type product
- South\_Africa

Taguatinga

Logged in as sen2agni | DB version 2.0.1

- Essas ferramentas se tornam úteis para enfrentar problemas sociais, como o combate a pobreza e a fome mundial;
- As crises internacionais de fornecimento de alimentos mostraram ainda como a produção agrícola e seus mercados estão globalmente conectados e exibem uma ampla variação geográfica e altas flutuações ao longo do tempo;
- É reconhecido atualmente que o impacto da volatilidade de mercado e a insegurança alimentar, não podem ser minimizados sem intervenções políticas baseadas em evidências científicas sólidas.
- Os dados de Observação da Terra podem contribuir para esse objetivo como uma fonte comprovada de informações transparentes, precisas e consistentes sobre a produtividade agrícola em escala global e regional.

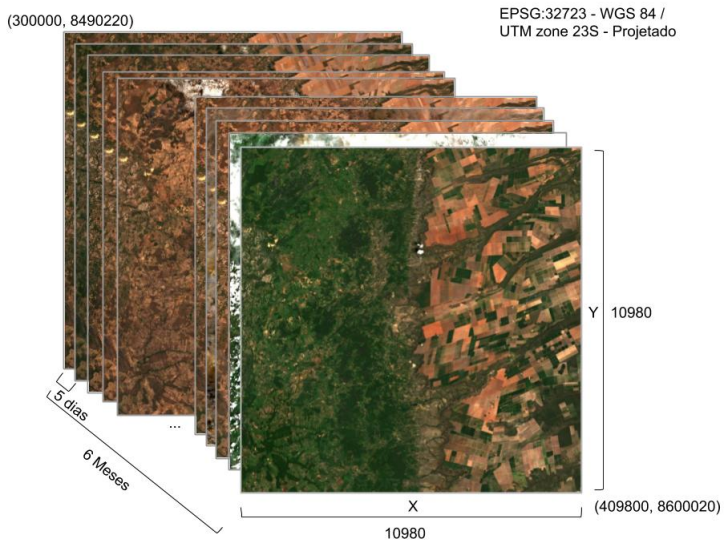


Figure: *Dados*

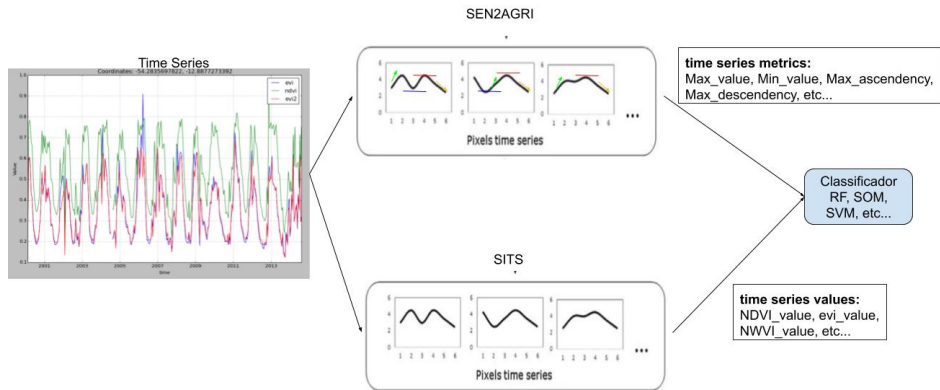


Figure: Métodos de extração de características

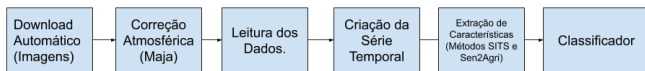


Figure: *Métodos de extração de características*

- O desenvolvimento foi realizado utilizando a linguagem Python.
- Programa para correção de imagens MAJA (MACCS-ATCOR Joint Algorithm).
- Bibliotecas principais: GDAL, NumPy, scikit-Learning
- Para executar o trabalho foi necessário utilizar a máquina- servidora (esensing-005)

# Matriz de Confusão e Acurácia

RF (Sen2Agri, max_depth = 5 ) / Acurácia = 0,965		
	Áreas Cultivadas	Áreas não Cultivadas
Áreas Cultivadas	91792	1761
Áreas não Cultivadas	3395	52853
	93553	56248

RF (Sen2Agri, max_depth = 15 ) = 15 / Acurácia = 0,998		
	Áreas Cultivadas	Áreas não Cultivadas
Áreas Cultivadas	93523	30
Áreas não Cultivadas	126	56122
	93553	56248

RF (Sen2Agri, max_depth = 25 ) / Acurácia = 1,0		
	Áreas Cultivadas	Áreas não Cultivadas
Áreas Cultivadas	93553	0
Áreas não Cultivadas	0	56248
	93553	56248

RF(SITS, max_depth = 5 ) / Acurácia = 0,975		
	Áreas Cultivadas	Áreas não Cultivadas
Áreas Cultivadas	90625	2928
Áreas não Cultivadas	788	55460
	93553	56248

RF (SITS, max_depth = 15) / Acurácia = 0,999		
	Áreas Cultivadas	Áreas não Cultivadas
Áreas Cultivadas	93553	21
Áreas não Cultivadas	0	56248
	93553	56248

RF(SITS, max_depth = 25) / Acurácia = 1,0		
	Áreas Cultivadas	Áreas não Cultivadas
Áreas Cultivadas	93553	0
Áreas não Cultivadas	0	56248
	93553	56248



# Máscara de áreas cultivadas

21-12-2017



MaskCrop SITS



MaskCrop Sen2AGRI



Obrigado.