

Instituto Nacional de Pesquisas Espaciais - INPE

Fabiana Zioti, Gabriel Sansigolo, Rafael Mariano

# **Detecção de Áreas Queimadas com o uso de Índices Espectrais**

São José dos Campos

2018



Fabiana Zioti, Gabriel Sansigolo, Rafael Mariano

# Detecção de Áreas Queimadas com o uso de Índices Espectrais

Trabalho Final Disciplina CAP-378 Tópicos em Observação da Terra. Prof: Antônio Miguel Vieira

São José dos Campos

2018



## Resumo

As queimadas são atividades que causam impactos negativos no âmbito social, ambiental e econômico. Essas consequências motiva diversas pesquisas relacionadas à ocorrência de incêndios e seus impactos em diversos biomas. O objetivo deste trabalho é realizar um estudo sobre os índices espectrais aliado a aplicação de métodos de aprendizado de máquina, com o propósito de verificar se uma perspectiva automática no processo de detecção de cicatrizes de queimadas possui um resultado similar ao método de detecção semiautomático, realizado pelos especialistas. Visando futuramente, aderir aos métodos automáticos no processo de detecção de queimadas.



## Lista de Figuras

2.1	Linha do tempo do Programa Landsat (Passado, Presente e o Futuro) . . .	13
2.2	Bandas dos sensores TIRS e OLI do satélite Landsat 8 . . . . .	14
2.3	Arquitetura Rede MLP . . . . .	17
2.4	SVM: Hiperplano . . . . .	18
2.5	K-Médias Exemplo . . . . .	19
3.1	Dados vetoriais - Programa Queimadas . . . . .	21
3.2	Metodologia do projeto . . . . .	22
3.3	Grade do ponto-órbita 221/067 do satélite Landsat 8 - OLI . . . . .	23
3.4	K-Médias da diferença do NDVI . . . . .	23
4.1	Mascara de saída do K-Médias 1 . . . . .	25
4.2	Mascara de saída do K-Médias 2 . . . . .	26
4.3	Mascara de saída do K-Médias 3 . . . . .	27
4.4	Comparação dos algoritmos 1 . . . . .	28
4.5	Comparação dos algoritmos 2 . . . . .	29





## Lista de Tabelas

4.1	Resultados do K-Médias NDVI . . . . .	26
4.2	Resultados do K-Médias NBRL . . . . .	27



# Sumário

1. <i>Introdução</i> . . . . .	11
2. <i>Referencial Teórico</i> . . . . .	13
2.1 Landsat . . . . .	13
2.2 Índices Espectrais . . . . .	15
2.3 Aprendizado de Máquina . . . . .	15
2.3.1 Redes Neurais Artificiais . . . . .	16
2.3.2 Árvore de Decisão . . . . .	17
2.3.3 Máquina Vetores de Suporte . . . . .	18
2.3.4 K-Médias . . . . .	19
3. <i>Metodologia</i> . . . . .	21
3.1 Área de Estudo . . . . .	22
3.2 Algoritmos de Aprendizado . . . . .	23
4. <i>Resultados</i> . . . . .	25
5. <i>Conclusões</i> . . . . .	31
<i>Referências</i> . . . . .	33



## Introdução

As queimadas podem ser iniciadas por razões naturais ou por iniciativas humanas. Elas liberam dióxido de carbono ( $CO_2$ ), uma composição química que é considerada uma das principais causas do aumento do efeito estufa (J. S. et al. (1995)). A queimada é o principal motivo da perda de vegetação no território brasileiro, apresentando consequências sociais, ambientais e econômicas. Isso motiva várias investigações relacionadas à ocorrência de incêndios e seu impacto no bioma do país.

A análise de incêndios tem duas vertentes de pesquisa: a primeira relaciona-se à detecção de focos, incêndios ativos e a segunda relacionada à detecção de cicatrizes de áreas queimadas, marcas deixadas pelo fogo após a sua extinção (Justice et al. (2002)).

Neste contexto, as imagens de satélite tornaram-se relevantes para monitorar eventos relacionados a incêndios. Os sensores a bordo dos satélites podem observar lugares inacessíveis e fazer uma varredura em grandes áreas em pouco tempo. O satélite Landsat 8, por exemplo, tem uma resolução temporal de 16 dias.

O programa Queimadas é um projeto desenvolvido pelo Instituto Nacional de Pesquisas Espaciais (INPE) com o objetivo de monitorar o risco e o impacto do fogo. Foram desenvolvidos diversos produtos como: coordenadas geográficas do foco, alertas por e-mail de ocorrências, risco de incêndio, estimativas de concentração de fumaça, mapeamento de áreas queimadas, entre outros. Eles permitem que os governos (federal, estadual e municipal) tomem decisões relacionadas a políticas públicas (INPE (2018)). Por este motivo a precisão dos dados é de suma importância.

O objetivo deste trabalho é realizar um estudo sobre os índices espectrais aliado a aplicação de métodos de aprendizado de máquina, com o propósito de verificar se uma perspectiva automática no processo de detecção de cicatrizes de queimadas possui um

resultado similar ao método de detecção semiautomático, realizado por especialistas. Visando futuramente, aderir a utilização de métodos automáticos ao processo de detecção de queimadas.

Usamos como base a pesquisa Pereira et al. (2016), que apontou os melhores índices espectrais para a região do Cerrado brasileiro, sendo eles: NDVI (Índice de Vegetação por Diferença Normalizada) e o MIRBI (Índice de Queimadura Infravermelha Média). Os dados que serão utilizados neste projeto são: (a) imagens do sensor OLI a bordo do satélite Landsat-8, órbita 221 ponto 067, para todo o ano de 2017; (b) dados vetoriais com as cicatrizes de queimadas geradas e validadas pelo Programa Queimadas do INPE.

## Referencial Teórico

### 2.1 Landsat

A série Landsat é um programa de satélites com mais de 40 anos de história. Este programa já lançou 8 instrumentos em órbita da Terra, como pode ser visualizado na Figura 2.1. Em Julho de 1972, aconteceu o lançamento de seu primeiro satélite o *Earth Resources Technology Satellite (ERTS)*, posteriormente chamado de Landsat 1. Ele foi um marco em sua época, sendo pioneiro no uso de plataformas espaciais para coleção sistemática de imagens terrestres. Hoje o programa Landsat possui um conjunto de quase meio século de imagens voltado a Observação da Terra, possibilitando que áreas como de Sensoriamento Remoto realizem pesquisas de monitoramento de uso e cobertura da terra, perturbações de causas naturais ou humanas, e afins (Williams et al. (2006); Loveland et al. (2008); USGS (2018)).

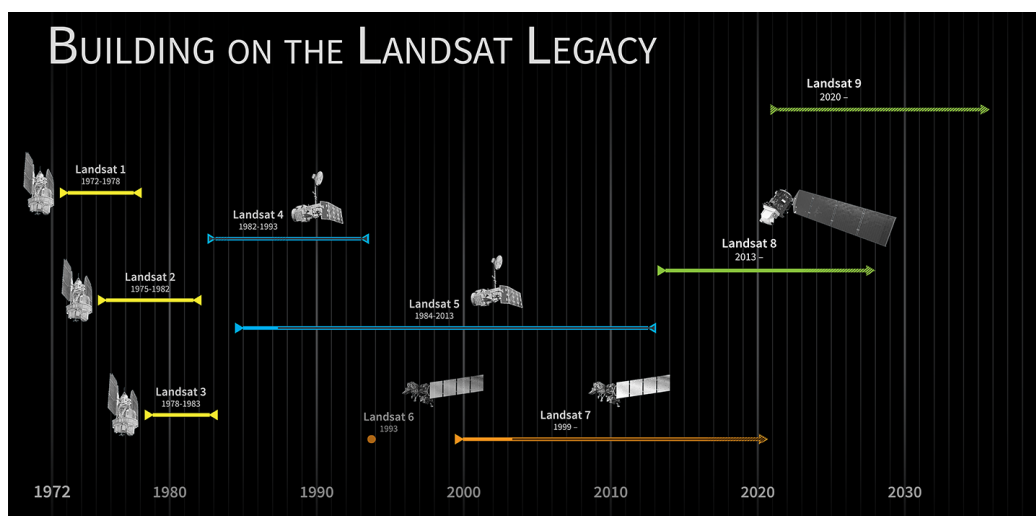


Figura 2.1: Linha do tempo do Programa Landsat (Passado, Presente e o Futuro)

Em 11 de fevereiro de 2013, foi lançado o mais recente satélite da série, o Landsat 8, na base da força aérea de Vandenberg, Califórnia. O Landsat 8 possui 2 sensores a bordo: *Thermal Infrared Sensor (TIRS)* e o *Operational Land Imager (OLI)*. Esses sensores, como é mostrado na Figura 2.2 contém 9 bandas espectrais, que captura informações no comprimento de onda do visível (bandas 1, 2 e 3), no vermelho (banda 4), infravermelho próximo (banda 5), SWIR (banda 6), SWIR 2 (banda 7), Pancromática (banda 8) e a Cirrus (banda 9). O Landsat 8 possui uma resolução temporal de 16 dias, e as faixas tem uma resolução espacial de 30 metros, com exceção da Pancromática com 15 metros. Este satélite é um resultado da parceria entre a *National Aeronautics and Space Administration (NASA)* e o *Department of the Interior U.S. Geological Survey (USGS)* (NASA (2018a); Roy et al. (2014)).

Landsat 8	
Band description (30 m native resolution unless otherwise denoted)	Wavelength ( $\mu\text{m}$ )
Band 1 – blue	0.43–0.45
Band 2 – blue	0.45–0.51
Band 3 – green	0.53–0.59
Band 4 – red	0.64–0.67
Band 5 – near infrared	0.85–0.88
Band 6 – shortwave infrared	1.57–1.65
Band 7 – shortwave infrared	2.11–2.29
Band 8 – panchromatic (15 m)	0.50–0.68
Band 9 – cirrus	1.36–1.38
Band 10 – thermal Infrared (100 m)	10.60–11.19
Band 11 – thermal Infrared (100 m)	11.50–12.51

Figura 2.2: Bandas dos sensores TIRS e OLI do satélite Landsat 8 (Adaptado em [5]).

Foi anunciado em 16 de Abril de 2015, o planejamento de lançamento para o próximo satélite do programa, o Landsat 9. A NASA e a USGS trabalham para terem um lançamento no ano de 2020. Segundo NASA (2018c), este satélite irá replicar em grande parte o seu irmão mais novo, o Landsat 8, e levará dois tipos diferentes de instrumentos, um para capturar comprimentos de ondas na faixa do visível, infravermelho próximo e infravermelho de ondas curtas e o segundo para medir radiação térmica infravermelha (NASA (2018b)).



## 2.2 Índices Espectrais

Índices espectrais são a combinação da refletância espectral de dois ou mais comprimentos de onda para destacar uma característica relativa de interesse, como vegetação, características artificiais, hídricas e geológicas. Essas combinações têm a vantagem de permitir uma comparação quantitativa Guyot e Gu (1994).

Neste estudo, será usado dois índices espectrais (NDVI e MIRBI) que já foram previamente calculados. Segundo Pereira et al. (2016), esses índices são os melhores para classificar áreas queimadas na região do Cerrado brasileiro. Logo abaixo, segue as fórmulas de cada índice espectral.

A fórmula do Índice de Vegetação por Diferença Normalizada (NDVI) é o resultado da seguinte equação Rouse Jr et al. (1974).

$$NDVI = \frac{(NIR + RED)}{(NIR - RED)} \quad (2.1)$$

A fórmula do índice Mid-Infrared Burn (MIRBI) é o resultado da seguinte equação Trigg e Flasse (2001).

$$MIRBI = (10 * SWIR2) - (9.8 * SWIR1) + 2 \quad (2.2)$$

## 2.3 Aprendizado de Máquina

Samuel (1959) descreveu Aprendizado de Máquina (AM) como um subcampo da Inteligência Artificial destinado a ensinar as máquinas a habilidade de pensar, sem serem programadas de forma explícita. Portanto um dos principais objetivos do AM é encontrar funções ou hipóteses, com base em um conjunto de dados, que podem ser usadas para prever ou classificar um novo exemplo.

Segundo Mitchell (1997), no aprendizado supervisionado temos um conjunto de dados de entrada com as respectivas respostas de cada exemplo. Cada exemplo de treinamento é um par na forma:

$$(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)}) \quad (2.3)$$

onde  $x^{(i)}, y^{(i)} \in \mathbb{R}^n$ . O objetivo neste caso é encontrar uma fronteira de decisão que

separe as classes do problema. Em contraste, no Aprendizado não Supervisionado os dados de entrada não possuem respostas associadas a eles. Cada exemplo de treinamento é definido por [2.4], neste caso o algoritmo os agrupa com base em algum critério de similaridade:

$$x^{(1)}, x^{(2)}, \dots, x^{(m)} \quad (2.4)$$

### 2.3.1 Redes Neurais Artificiais

As Redes Neurais Artificiais simulam a maneira de transmissão e aprendizagem de informação dos neurônios presentes no cérebro humano. RNA's são definidas como uma técnica de otimização onde se busca por uma hipótese de uma função recorrendo a otimização da mesma, ou seja, minimizar (ou otimizar) uma função.

Para simular as sinapses biológicas executadas pelo cérebro, os pesquisadores desenvolveram um modelo de neurônio artificial.

Dado  $x^{(m)}, y^{(m)}$  e um vetor de peso  $\theta \in \mathbb{R}^n$ , a operação [2.5] é executada pelo neurônio:

$$z = \sum_j^n x_j^i \theta_j \quad (2.5)$$

Com  $z \in \mathbb{R}^n$ . É aplicado sobre  $z$  um função de ativação, e assim temos a saída do neurônio.

$$h_\theta(x) = g(z) \quad (2.6)$$

Os valores de  $\theta$  determina o grau de importância de cada entrada do neurônio. Portanto o aprendizado consiste em encontrar os melhores valores para  $\theta$ , com base em um conjunto de entrada, de tal forma que a sua hipótese seja a mais próxima da saída desejada e diante disso classifique corretamente os exemplos de treinamento.

Existem diversos modelos de ANN: Perceptron, Auto Encoder, Convolution, Kohonen são alguns exemplos. Neste trabalho utilizamos a Rede Neural Perceptron Multicamadas (MLP). A rede MLP é um modelo composto por múltiplas camadas de neurônios artificiais totalmente conectadas, denominadas: Camada de Entrada, Camada Intermediária e Camada de Saída. Um exemplo de sua arquitetura pode ser visualizada em 2.3

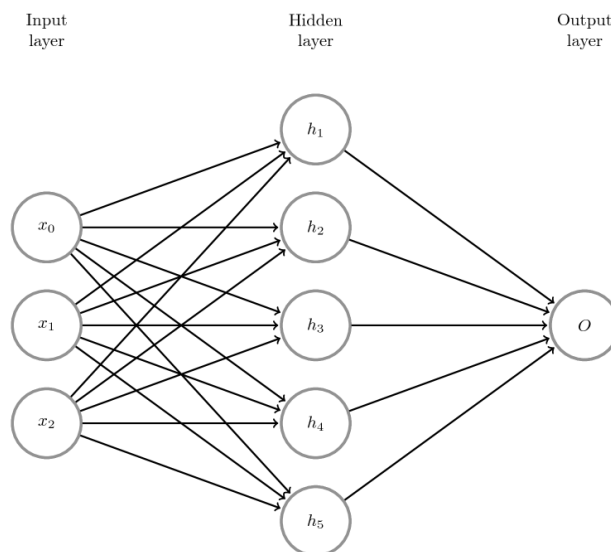


Figura 2.3: Arquitetura Rede MLP (MaviccPRP (2017))

A camada de entrada é responsável por receber os valores  $x^{(1)}, x^{(2)}, \dots, x^{(m)}$ . Dessa forma, essa camada define o número de entradas na RNA. A camada intermediária, também é denominada camada escondida, para maioria dos problemas uma camada é suficiente. Entretanto alguns pesquisadores trabalham com um número maior de camadas escondidas. A camada de saída projeta o número de classes que será classificado pela RNA. É a saída da rede, definida por  $h(x)$ .

Para que a MLP consiga prever com exatidão a hipótese é necessário realizar seu treinamento. Por se tratar de um algoritmo de aprendizado supervisionado, devemos apresentar o conjunto  $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})$ .

### 2.3.2 Árvore de Decisão

A árvore de decisão é uma estrutura de dados que pode ser representada como um conjunto de regras. As árvores classificam os dados percorrendo da raiz até chegar ao nó folha, este que representa as classes do conjunto. Ou seja, a árvore possui duas estruturas de nó:

- Nó folha: Corresponde a uma classe.
- Nó não-folha: Contém um teste de algum atributo dos dados. Suas ramificações correspondem a um dos valores possíveis para esse atributo.

A classificação é realizada começando no nó raiz, é feito a aplicação dos teste, de forma a averiguar qual ramificação será seguida: a sub-árvore à esquerda ou sub-árvore à direita. É executada a caminhada na árvore, até que um nó folha seja encontrado. Como resultado, o novo exemplo será definido pela classe do nó folha.

A árvores de decisão possuem diversos algoritmos de aprendizado. De acordo com Mitchell (1997) os amplamente utilizados são ID3, ASSISTANT, e C4.

Árvores de decisão são comumente usadas como ferramenta visual e analítica de apoio à decisão, por que com ela é possível se calcular as probabilidades condicionais.

### 2.3.3 Máquina Vetores de Suporte

Em uma classificação binária linear com dados  $x, y$ , com  $y \in -1, 1$  e os parâmetros  $\theta$  e  $b$  (*bias*), o objetivo é encontrar o hiperplano que separa os dados em duas classes distintas, o classificador Máquina Vetores de Suporte, do inglês *Support Vector Machine (SVM)*, pode ser definido como NG (2012):

$$z = \theta^T x + b \quad (2.7)$$

$$h_{\theta,b}(x) = g(z) \quad (2.8)$$

Onde  $g(z) = 1$  se  $z \geq 0$  ou  $g(z) = -1$  caso contrario.

Um conceito importante no SVM é a margem. Ela é a distância entre o hiperplano e o primeiro ponto de cada classe, apresentado na Figura 2.4.

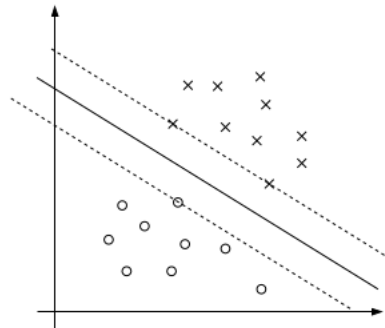


Figura 2.4: SVM: Hiperplano (NG (2012))

O objetivo neste caso é encontrar uma fronteira de decisão que maximize a margem,

gerando como resultado um conjunto de previsões com um bom ajuste. Em suma, o SVM é um classificador que separa os dados positivos e negativos usando um hiperplano, com intervalo de lacuna (margem).

### 2.3.4 K-Médias

O principal objetivo do algoritmo de aprendizado não supervisionado K-Médias é particionar um conjunto de dados  $x^m$  em K números de agrupamentos.

O algoritmo K-Médias compreende dois passos principais. Após o agrupamento por centroides aleatórios  $\mu_1, \mu_2, \dots, \mu_k \in \mathbb{R}^n$ , o primeiro passo é atribuir todos os exemplos  $x^i$  em um grupo baseado no centroide mais próximo [2.9], o segundo passo é atualizar o centroide com base na média do seu grupo (Usman (2013)).

$$c^{(i)} = \operatorname{argmin}_k \|x^{(i)} - \mu_k\|^2 \quad (2.9)$$

Os passos devem ser repetidos até uma convergência, ou seja, minimizar a função custo, definida por [2.10]:

$$J = \frac{1}{m} \sum_{i=1}^m \|x^{(i)} - \mu_K\|^2 \quad (2.10)$$

A Figura 2.5 apresenta um exemplo visual do particionamento de um espaço com o K-Médias.

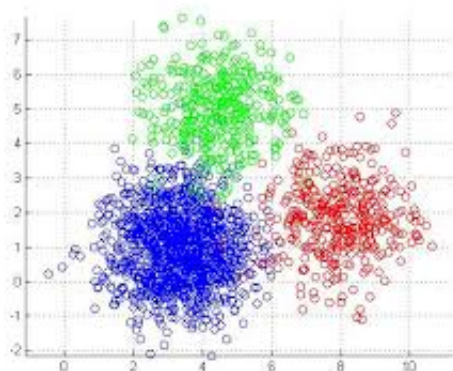


Figura 2.5: K-Médias Exemplo (Chen (2017))



## Metodologia

Nesta seção, apresentamos a metodologia aplicada no projeto. Atualmente o processo de detecção de cicatrizes de queimadas, no contexto do Programa Queimadas do INPE, é realizado de maneira semi-automática. Primeiramente é gerado os índices espectrais NDVI e NBRL, no próximo passo é realizado o cálculo da diferença dos índices com uma imagem do tempo anterior, para destacar áreas que houveram mudanças na vegetação e posteriormente é aplicando um limiar, que é definido por um especialista. Este limiar define os grupos de áreas queimadas e não queimadas. Por fim, é gerado uma imagem binária com as possíveis regiões de queimadas, que a partir dela é feito uma vetorização, Figura 3.1. Os dados vetoriais, são verificados pelo especialista confirmando se são ou não áreas de queimada.

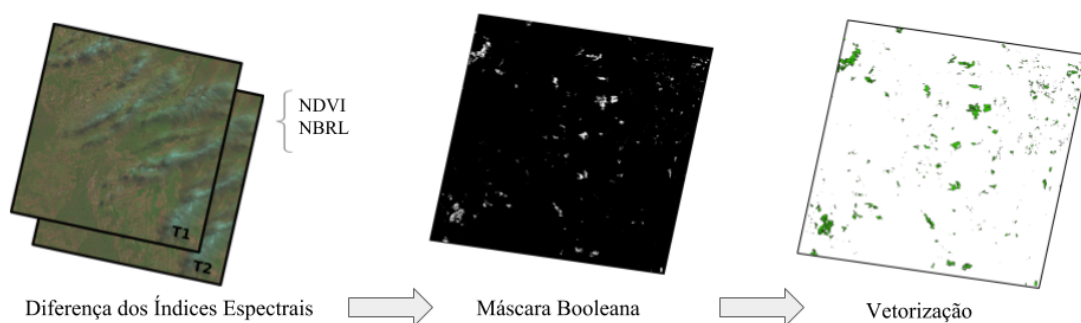


Figura 3.1: Dados vetoriais - Programa Queimadas)

Para o treinamento dos métodos de aprendizado supervisionado, utilizamos os dados vetoriais com as cicatrizes de queimadas geradas e validadas pelo Programa Queimadas. Os dados possuem atributos como: NDVI, a diferença do NDVI, NBRL, a diferença do NBRL, e as medianas das bandas [1, 7] do Landsat 8. A diferença dos índices espectrais são

importantes pois mostram um indicativo de mudança na imagem, o que pode ser causado pelas queimadas. Para isso, é realizado a subtração de uma cena de uma data atual com a cena de uma data anterior, Figura 3.2.

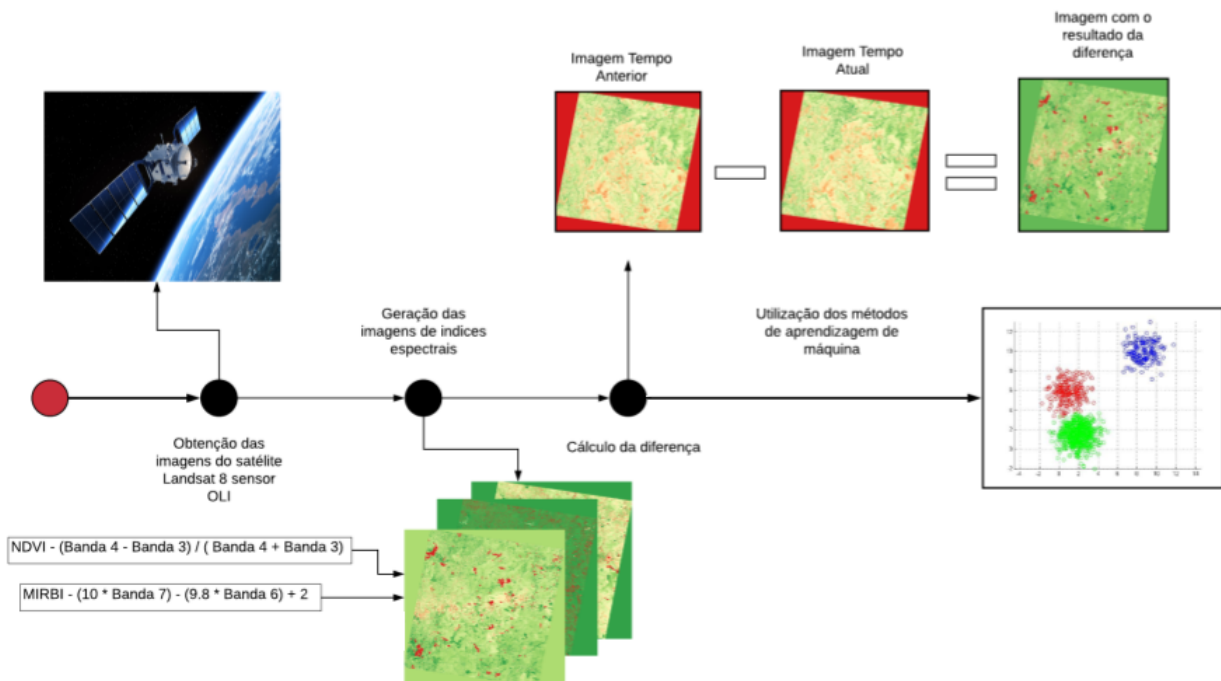


Figura 3.2: Metodologia do projeto

Os dados para análise é um arquivo no formato shapefile, no período entre 05 de Maio de 2017 á 12 de Outubro de 2017.

### 3.1 Área de Estudo

Para a área de estudo, foi escolhido o ponto órbita 221/067 do satélite Landsat 8 sensor OLI, como é demonstrado na Figura 3.3. Essa região de estudos é importante, pois inclui a área do Jalapão que é seriamente ameaçada por queimadas florestais. O Jalapão possui a maior área de conservação ambiental do Bioma Cerrado, protegida pelo Ministério do Meio Ambiente do Brasil através do Projeto Cerrado-Jalapão <sup>1</sup>. Este projeto possui diversas atividades que visam preservar e controlar os eventos de incêndio no Bioma Cerrado.

<sup>1</sup> <http://cerradojalapao.mma.gov.br/projeto>



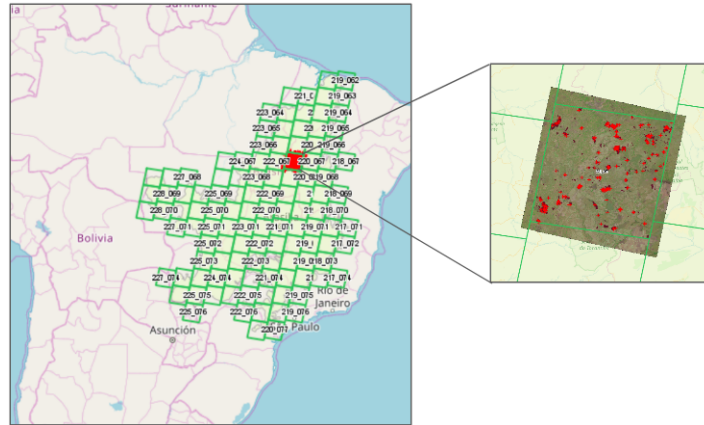


Figura 3.3: Grade do ponto-órbita 221/067 do satélite Landsat 8 - OLI

### 3.2 Algoritmos de Aprendizado

Empregamos a perspectiva automática no processo de detecção de cicatrizes de queimadas em duas situações. A primeira utilizamos o método K-Médias com as imagens da diferença do índices espectrais, individualmente, como entrada. Foram utilizadas as cenas do Landsat 8 nas datas: 06/06/2017 e 22/06/2017, 09/08/2017 e 25/09/2017, 10/09/2017 e 26/09/2017.

O algoritmo foi executado com valores de  $k$  no intervalo  $[2, 8]$ . Após análise visual, concluiu-se que  $K = 3$  alcança o melhor resultado, destacando as áreas queimadas. O resultado do algoritmo é apresentado na Figura 3.4.

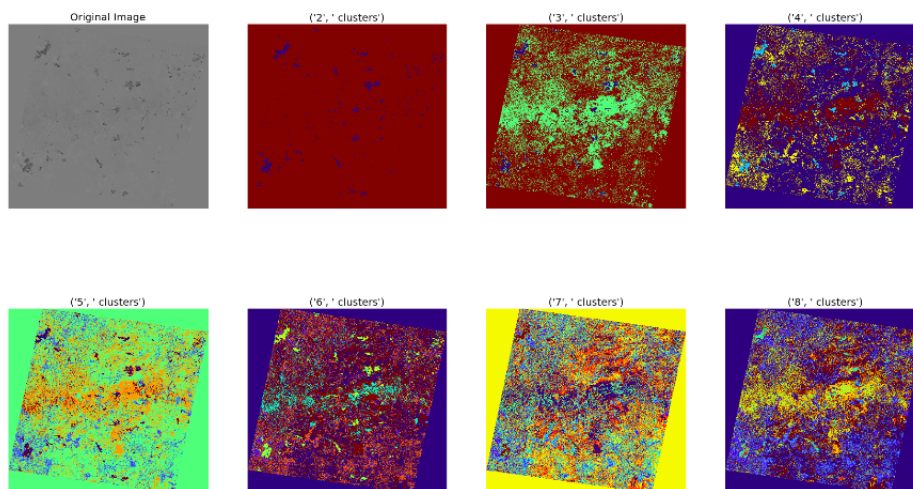


Figura 3.4: K-Médias da diferença do NDVI

Verificou-se também que o agrupamento da diferença do NDVI obteve os melhores resultados, que serão apresentados no próximo capítulo. Nesta situação o algoritmo atuou como o limiar separando as áreas que sofreram mudanças. A segunda situação é análoga a verificação dos dados vetoriais, realizada pelo especialista. O intuito foi investigar se a abordagem automática possui um resultado similar a verificação manual.

Neste contexto utilizamos os algoritmos SVM, MLP e a Árvore de Decisão. Para realizar o treinamento destes métodos, utilizamos os atributos NDVI, Diferença do NDVI, NBRL, Diferença do NBRL, e as medianas das bandas [2, 5] dos dados vetoriais. Os dados vetoriais possui também uma coluna verifica, esta coluna indica se a área em questão é uma cicatriz de queimadas ou não.

Foram realizados dois teste, o primeiro utilizando os índices espectrais e suas diferenças. No segundo teste, foi adicionado aos dados de entrada as medianas das bandas [2, 5]. Os resultados são apresentados na próxima seção.

## Resultados

Como apresentado na metodologia, a abordagem via K-Médias apresentou bons resultados para o destaque de áreas queimadas. São apresentados os resultados do K-Médias, com as entradas a diferença do índice espectral NDVI e a diferença do índice espectral NBRL. Com o propósito de melhorar a visualização dos resultados será apresentado apenas imagens da diferença do NDVI.

Na Figura 4.1, é apresentado os resultados do algoritmo K-Médias no processo de detecção de áreas queimadas. A máscara de saída está em cores preto e branco, e as áreas queimadas, verificadas e validadas pelo programa Queimadas, estão com contornos em cor vermelho. Este é o resultado do processamento descrito na metodologia com as cenas entre o dia 10 de Setembro de 2017 e 26 de Setembro de 2017.

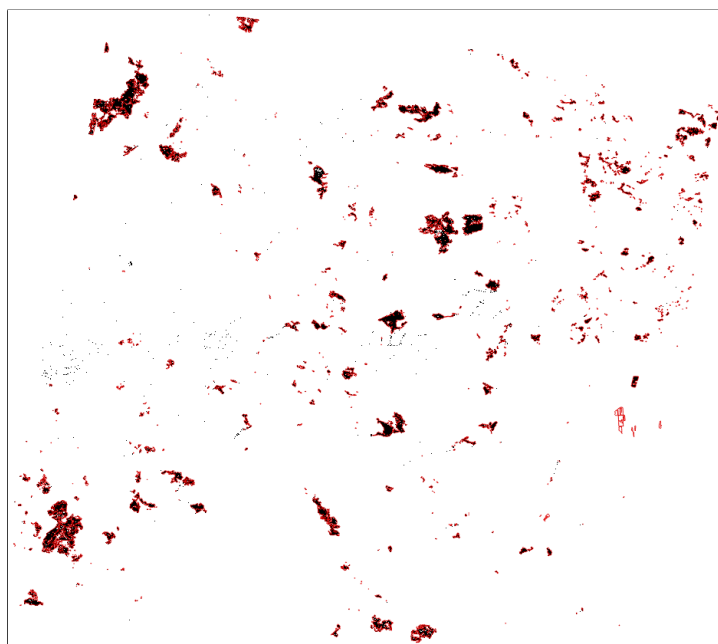


Figura 4.1: Mascara de saída do K-Médias 1

Na Figura 4.2, é apresentado o resultado do processamento com as cenas entre o dia 09 de Agosto de 2017 e 25 de Outubro de 2017.

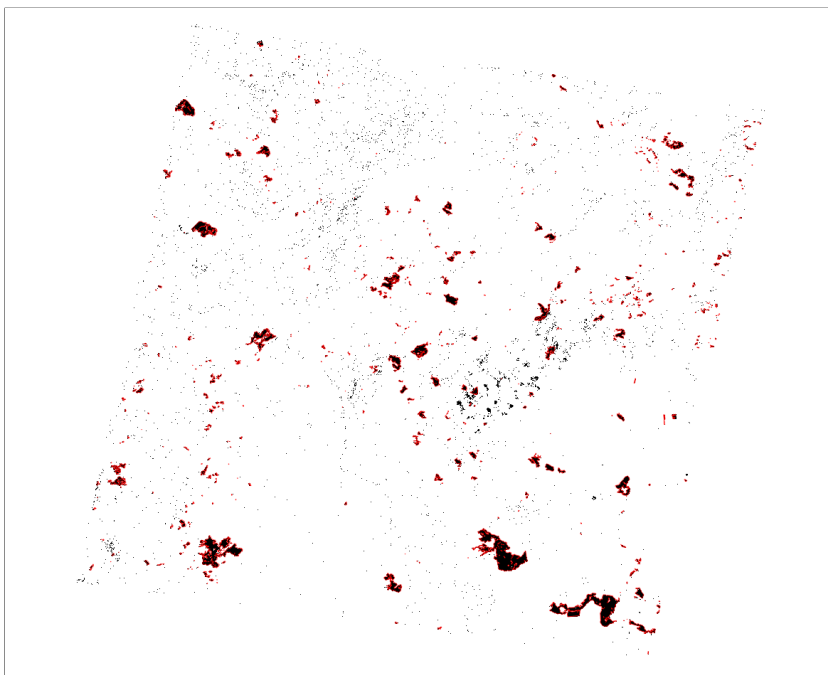


Figura 4.2: Mascara de saída do K-Médias 2

Na Tabela 4.1 é apresentado a porcentagem de áreas queimadas e não queimadas sobre a máscara de cicatrizes.

	Porcentagem de Queimada	Porcentagem de Não-Queimada
<b>Mascara</b>	98.31	1.69
<b>Dados Validados</b>	98.19	1.81

Tabela 4.1 - Resultados do K-Médias NDVI

Com a abordagem K-Médias alcançamos 99,87% de precisão na data 10/09/2017 e 26/09/2017, em relação à máscara verificada e validada do Programa Queimadas. Entretanto o K-Médias apresentou resultados insatisfatórios, como mostra a Figura 4.3, para a data entre o dia 06 de Junho de 2017 e 22 de Junho de 2017.

Na Tabela 4.2 é apresentando o resultado das porcentagens de áreas queimadas e não queimadas sobre o índice espectral NBRL.

A abordagem K-Médias sobre a diferença de NDVI, apresentou resultados satisfatórios na área analisada, principalmente no processamento das datas entre 10 de Setembro de 2017 e 26 de Setembro de 2017. O resultado satisfatório nesta data se deve a motivos

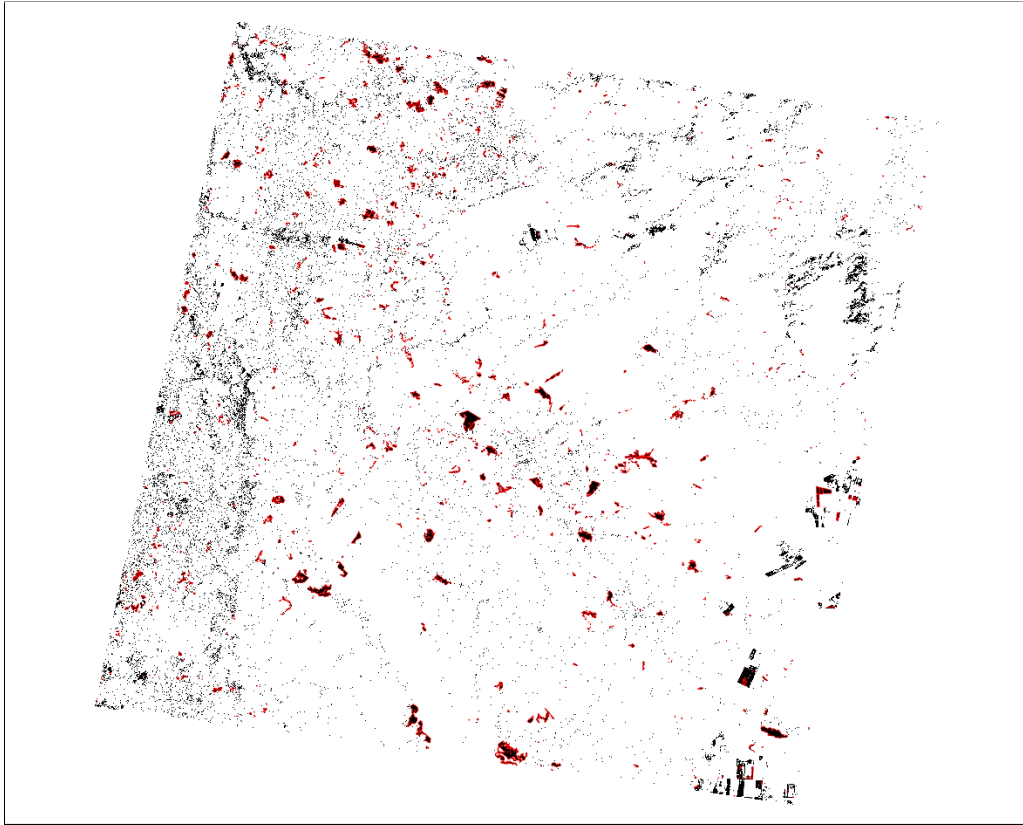


Figura 4.3: Mascara de saída do K-Médias 3

	Porcentagem de Queimada	Porcentagem de Não-Queimada
<b>Mascara</b>	68.55	31.45
<b>Dados Validados</b>	98.19	1.81

Tabela 4.2 - Resultados do K-Médias NBRL

que desconhecemos. Formulamos hipóteses, como estação do ano, período chuvoso, etc. Porém esta questão ainda deve ser analisado com mais cautela e não situa-se no escopo deste projeto.

A seguir são apresentados os resultados dos algoritmos de aprendizagem supervisionado. Foram realizados dois teste. No primeiro os algoritmos receberam como entrada NDVI, Diferença do NDVI, NBRL, Diferença do NBRL. O conjunto de entrada neste caso ficou definido como  $x^{(i)} \in \mathbb{R}^4, y^{(i)} \in 0, 1$ . Para avaliar a performance dos algoritmos foi utilizada a função *Cross-validation* da biblioteca *scikit-learn*. A avaliação dos dados não deve ser feita utilizando o conjunto de treinamento. Uma das metologias de avaliação consiste em dividir os dados em dois conjuntos: teste e treinamento. A figura tal apresenta o resultados do *Cross-validation* para o conjunto de dados divididos (treinamento e teste), calculando

a pontuação 7 vezes consecutivas (com diferentes divisões de cada vez) para os modelos SVM, MLP e Árvore de Decisão.

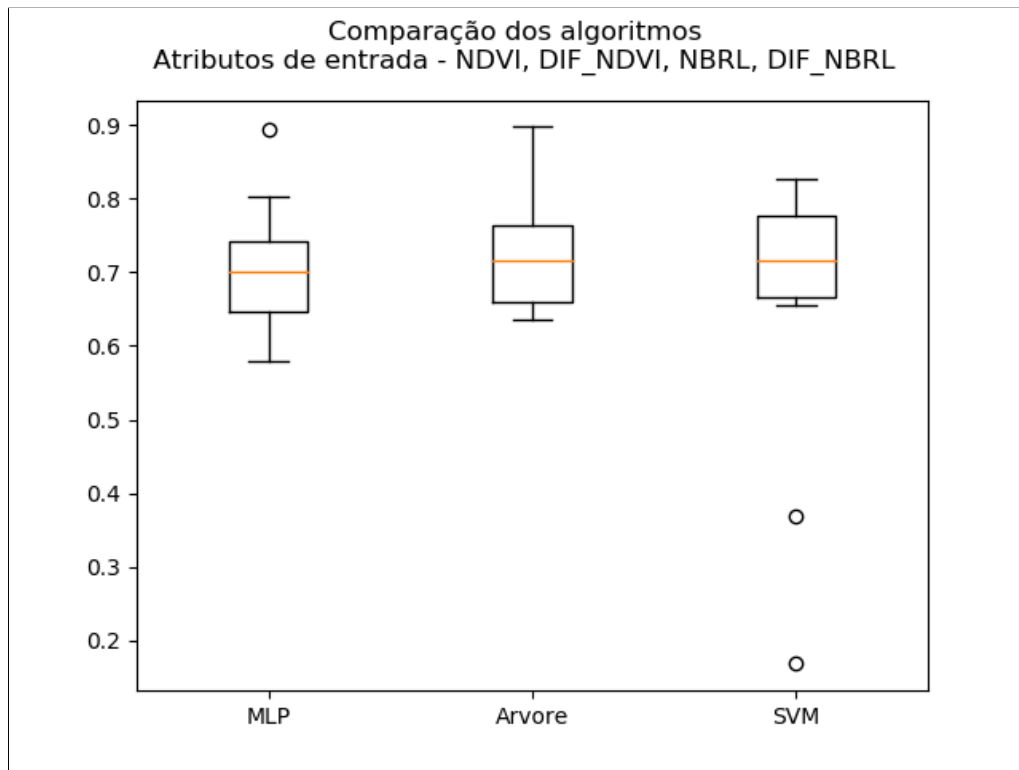


Figura 4.4: Comparaç o dos algoritmos 1

No segundo teste os algoritmos receberam como entrada os atributos dos  ndices espectrais NDVI, Diferenç a do NDVI, NBRL, Diferenç a do NBRL, e as medianas das bandas [2, 5]. O conjunto de entrada neste caso ficou definido como  $x^{(i)} \in \mathbb{R}^8, y^{(i)} \in \{0, 1\}$ . A figura 4.5 apresenta o resultados do *Cross-validation* para o conjunto de dados divididos (treinamento e teste), calculando a pontuaç o 7 vezes consecutivas (com diferentes divis es de cada vez) para os modelos SVM, MLP e  rvore de Decis o.

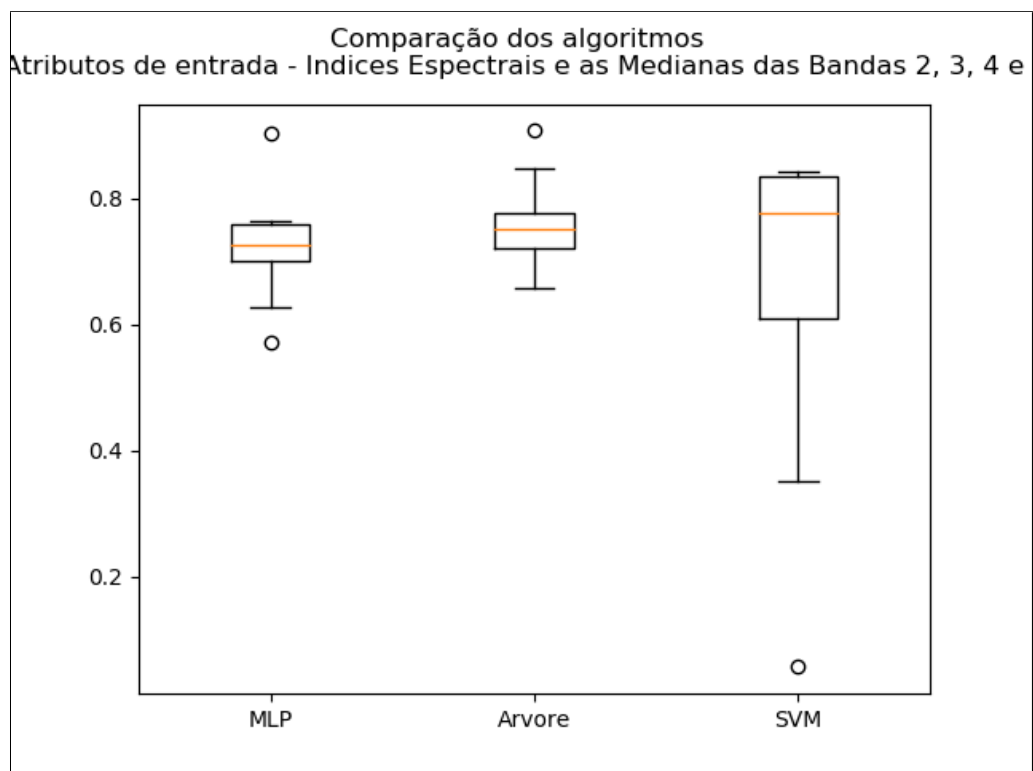


Figura 4.5: Comparação dos algoritmos 2





## Conclusões

Este trabalho apresentou um estudo sobre os índices espectrais aliado a aplicação de métodos de aprendizado de máquina. O intuito foi desenvolver um estudo sobre os métodos de aprendizado de máquina SVM, MLP, Árvore de Decisão e K-Médias, para investigar se estas técnicas apresentam resultados similar ao método de detecção semiautomático, realizado pelos especialistas.

Por se tratar de um estudo preliminar, não buscamos substituir o processo realizado atualmente. Como apresentado durante o documento, os métodos de aprendizado desenvolvidos neste trabalho apresentaram resultados preliminares satisfatórios.

O método K-Médias apresentou bons resultados atuando como limiar, utilizando como entrada a diferença do NDVI. Porém, quando apresentado sobre um outro índice espectral, como o NBRL, o resultado ficou baixo. Dito isso, como foi utilizado uma base de dados pequena não é possível afirmar se o NDVI é melhor que o NBRL, pois nem sempre os dados são os mesmos em cada imagem.

Com um estudo mais profundo destes métodos em conjunto com uma base de dados maior, pode vir a ser que a detecção de cicatrizes de queimadas utilizando métodos automáticos, apresentem resultados consideráveis e mais confiáveis.

Propomos como trabalho futuro a investigação de outras técnicas, entre elas a análise das séries temporais dos dados vetoriais, afim de eliminar os falsos positivos gerados por áreas agrícolas, analisando o seu comportamento espectral ao longo do tempo.



## Referências Bibliográficas

- Chen M., , 2017 <https://fr.mathworks.com/matlabcentral/fileexchange/24616-kmeans-clustering> <https://mavicccprp.github.io/a-neural-network-from-scratch-in-just-a-few-1>
- Guyot G., Gu X.-F., Effect of radiometric corrections on NDVI-determined from SPOT-HRV and Landsat-TM data, *Remote Sensing of Environment*, 1994, vol. 49, p. 169
- INPE, 2018 Portal do Monitoramento de Queimadas e Incêndios <http://www.inpe.br/queimadas>
- J. S. L., W. R. Cofer I., D. R. Cahoon J., E. L. W., , 1995 Technical report Biomass Burning: A Driver for Global Change!
- Justice C.O.and Giglio B., Korontzi S., Owens J., Morisette J., Roy D., Descloitres J., Alleaume S., Petitcolin F., Kaufman Y., The MODIS fire products, *Remote Sensing of Environment*, 2002, pp 244–262
- Loveland T. R., Cochrane M. A., Henebry G. M., Landsat still contributing to environmental research, *Trends in ecology & evolution*, 2008, vol. 23, p. 182
- MavicccPRP, 2017 A Neural Network from scratch in just a few Lines of Python Code <https://mavicccprp.github.io/a-neural-network-from-scratch-in-just-a-few-lines-of-python-code/>
- Mitchell T. M., *Machine learning*. McGraw-Hill, 1997
- NASA, 2018a Landsat 8 <https://landsat.gsfc.nasa.gov/landsat-data-continuity-mission/>

- NASA, 2018b Landsat 9 <https://landsat.gsfc.nasa.gov/landsat-9/>
- NASA, 2018c Landsat 9 Overview <https://landsat.gsfc.nasa.gov/landsat-9/landsat-9-overview/>
- NG A., , 2012 Cs 229 lecture notes: Support vector machines.
- Pereira A., Rodrigues Teixeira F., Libonati R., Melchiori E., L., Carvalho L., Evaluation of Spectral Indices for Burned Area Identification in Cerrado using Landsat TM Data, 2016, vol. 8, p. 1665
- Rouse Jr J., Haas R., Schell J., Deering D., Monitoring vegetation systems in the Great Plains with ERTS, 1974
- Roy D. P., Wulder M., Loveland T. R., Woodcock C., Allen R., Anderson M., Helder D., Irons J., Johnson D., Kennedy R., et al., Landsat-8: Science and product vision for terrestrial global change research, Remote sensing of Environment, 2014, vol. 145, p. 154
- Samuel A. L., Some Studies in Machine Learning Using the Game of Checkers, IBM Journal of Research and Development, 1959, vol. 3, p. 210
- Trigg S., Flasse S., An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah, International Journal of Remote Sensing, 2001, vol. 22, p. 2641
- USGS, 2018 Landsat Project Description <https://landsat.usgs.gov/landsat-project-description>
- Usman B., Satellite Imagery Land Cover Classification using K-Means Clustering Algorithm: Computer Vision for Environmental Information Extraction, 2013, pp 18671–18675
- Williams D. L., Goward S., Arvidson T., Landsat, Photogrammetric Engineering & Remote Sensing, 2006, vol. 72, p. 1171