# Detection of Planting Failures in Eucalyptus Stands Using Spectral Indices and Machine Learning on Sentinel-2 Imagery

**João Gabriel V. Camargo[1], Luis Guilherme V. Camargo[1], Giovani L. Dos Reis[2], Carlos R. dos Santos Junior[1]**

[1]Instituto Federal de Mato Grosso do Sul (IFMS)
Três Lagoas – MS – Brasil

[2]Universidade de São Paulo (USP)

{joao.camargo3, luis.camargo2}@estudante.ifms.edu.br

giovanireis@usp.br, carlos.santos@ifms.edu.br

***Abstract.*** *Eucalyptus cultivation plays a central economic role in regions such as the Mato Grosso do Sul pulp valley, where large-scale plantations drive regional development. Early detection of planting failures is critical for reducing losses and supporting operational decision-making. This study investigates whether supervised Machine Learning algorithms outperform a dynamic NDVI threshold baseline in detecting planting failures when fed exclusively with simple spectral features. Sentinel-2 images acquired in April 2024 were processed in QGIS, segmented into 30m $\times$ 30m cells, and manually labeled as normal or failure. Statistical features derived from NDVI and spectral bands B04 and B08 were organized into four attribute groups. Random Forest, SVM, Logistic Regression, and MLP models were trained on a development set (506 samples) and evaluated on an independent test set (545 samples) from a different stand. The NDVI baseline achieved the highest F1-score (0.7870), outperforming all ML models. This result shows that simple spectral features limit the advantage of ML models over a basic threshold rule, contributing to guide methodological choices in detection problems with restricted feature spaces. Among ML algorithms, MLP yielded the best isolated result (F1 = 0.7788). Future work will incorporate spatial context and multitemporal analysis to enrich feature representation and better exploit ML capabilities.*

***Resumo.*** *O cultivo de eucalipto tem papel econômico central em regiões como o Vale da Celulose de Mato Grosso do Sul, onde grandes plantações impulsionam o desenvolvimento regional. A detecção precoce de falhas nesses plantios é essencial para reduzir perdas e apoiar a tomada de decisão operacional. Este trabalho investiga se algoritmos de Aprendizado de Máquina supervisionados superam um baseline de limiar de NDVI na detecção de falhas de plantio quando alimentados exclusivamente por atributos espectrais simples. Imagens Sentinel-2 coletadas em abril de 2024 foram processadas no QGIS, segmentadas em células de 30m $\times$ 30m e rotuladas manualmente como normal ou falha. Estatísticas derivadas do NDVI e das bandas B04 e B08 foram organizadas em quatro grupos de atributos. Os modelos Random Forest, SVM, Regressão Logística e MLP foram treinados com um conjunto de desenvolvimento (506 amostras) e avaliados em um conjunto de teste independente (545 amostras)*

*de um talhão distinto. O Baseline NDVI obteve o melhor desempenho (F1 = 0,7870), superando todos os modelos de ML. Esse resultado mostra que atributos espectrais simples limitam o ganho dos modelos de ML em relação a uma regra de limiar básica, contribuindo para orientar escolhas metodológicas em problemas de detecção com espaço de atributos restrito. Entre os algoritmos de ML, o MLP obteve o melhor resultado isolado (F1 = 0,7788). Como trabalho futuro, pretende-se incorporar contexto espacial e análise multitemporal para enriquecer a representação de atributos e explorar melhor o potencial dos algoritmos de ML.*

## 1. Introdução

A agricultura brasileira ocupa posição estratégica na economia nacional, sendo o país um dos maiores produtores e exportadores de produtos agropecuários do mundo. Nesse contexto, o monitoramento contínuo das culturas tornou-se uma atividade essencial para identificar problemas de forma antecipada e auxiliar na tomada de decisão [Massruhá et al. 2020]. Fatores como déficit hídrico, pragas e deficiências nutricionais podem causar estresse nas plantas, comprometendo seu desenvolvimento e gerando perdas econômicas significativas.

Entre as culturas de relevância econômica regional, o eucalipto se destaca no estado de Mato Grosso do Sul, especialmente no chamado Vale da Celulose, onde multinacionais do setor de papel e celulose mantêm extensas plantações que movimentam a economia local. A detecção precoce de falhas nesses plantios — regiões onde o desenvolvimento vegetativo está comprometido ou inexistente — é crítica para que os produtores possam realizar replantios dentro da janela de tempo adequada, minimizando perdas de produtividade e custos operacionais.

Nos últimos anos, o sensoriamento remoto tem sido amplamente utilizado no monitoramento agrícola devido à disponibilidade de imagens de satélite de acesso público e à possibilidade de acompanhar grandes áreas de forma rápida e com baixo custo [Jensen 2009]. A partir dessas imagens, o índice NDVI (*Normalized Difference Vegetation Index*) pode ser calculado para avaliar o vigor vegetativo e identificar alterações associadas ao estresse das plantas [Huang et al. 2019, Leivas et al. 2022]. Abordagens baseadas em limiares de NDVI demonstraram resultados satisfatórios na delimitação de áreas com problemas em plantios de eucalipto [Galvão 2024], de forma que contribuem com novas abordagens metodológicas e se posiconam como referência comparativa (*baseline*).

Nesse contexto, este trabalho investiga a seguinte questão: *algoritmos de Aprendizado de Máquina supervisionados, quando alimentados exclusivamente por atributos espectrais simples, são capazes de superar um baseline de limiar de NDVI na detecção de falhas de plantio em eucalipto?* A hipótese investigada é que, quando os modelos utilizam apenas estatísticas simples do NDVI como entrada, algoritmos complexos de ML não apresentam vantagem significativa sobre uma regra de limiar. Para que os modelos de ML superem o baseline, seria necessário incluir informações adicionais, como o estado das células vizinhas e a variação temporal das imagens ao longo do ciclo de crescimento do eucalipto.

Para investigar essa hipótese, este trabalho propõe uma comparação sistemática

entre um baseline de limiar dinâmico de NDVI e quatro algoritmos de Aprendizado de Máquina amplamente consolidados na literatura — Random Forest, SVM, Regressão Logística e MLP [Reis et al. 2018] —, avaliados sobre atributos extraídos de imagens Sentinel-2 de plantios de eucalipto em Mato Grosso do Sul. Os conjuntos de desenvolvimento e teste foram extraídos de talhões distintos, garantindo avaliação independente e mais rigorosa da capacidade de generalização dos modelos. Os resultados obtidos contribuem para orientar escolhas metodológicas em problemas de detecção com espaço de atributos restrito, evidenciando os limites de abordagens que ignoram contexto espacial e temporal.

O restante do artigo está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados, a Seção 3 aborda os conceitos e técnicas utilizados, a Seção 4 descreve a metodologia, a Seção 5 apresenta e discute os resultados obtidos, e a Seção 6 apresenta as conclusões e trabalhos futuros.

## 2. Trabalhos Relacionados

O uso de sensoriamento remoto para monitoramento agrícola tem sido amplamente explorado por meio da análise de índices de vegetação extraídos de imagens orbitais [Jensen 2009]. [Huang et al. 2019] investigaram a utilização de séries temporais de NDVI e EVI para a identificação de fases fenológicas de culturas agrícolas, demonstrando que diferentes índices apresentam desempenhos distintos conforme o tipo de cultivo analisado e propondo um método de limiar dinâmico para detecção automatizada de eventos fenológicos. De forma complementar, [Leivas et al. 2022] avaliaram a dinâmica temporal dos índices NDVI, EVI, SAVI e IRECI utilizando imagens Sentinel-2A e MODIS, evidenciando o potencial dessas métricas para o acompanhamento do desenvolvimento vegetativo e a detecção de alterações na cobertura agrícola.

Na identificação de anomalias em plantios florestais, [Galvão 2024] utilizou imagens do satélite CBERS-4A e o índice NDVI para detectar falhas em plantios de eucalipto em fazendas do estado de São Paulo, comparando os padrões espectrais identificados nas imagens com registros fotográficos de campo. Esse trabalho serve como referência metodológica para o critério visual de rotulagem adotado no presente estudo, por estabelecer parâmetros concretos para distinguir regiões normais de regiões com falha a partir da análise do NDVI em talhões florestais. [Santos et al. 2020] empregaram sistemas aéreos não tripulados para monitorar o comportamento fenológico de espécies florestais, demonstrando a viabilidade do uso de imagens de alta resolução para acompanhamento da vegetação em escala local.

No que diz respeito à aplicação de algoritmos de Aprendizado de Máquina em dados de sensoriamento remoto florestal, [Reis et al. 2018] compararam métodos de regressão e algoritmos de ML — Random Forest, SVM e Redes Neurais Artificiais — para estimar atributos biofísicos de plantios de eucalipto a partir de imagens Landsat. O estudo mostrou que o Random Forest apresentou menor erro de estimativa em relação aos demais métodos, e que atributos espectrais derivados do NDVI e do infravermelho foram os mais relevantes nos modelos. Esses quatro algoritmos são amplamente utilizados como referência em problemas de classificação em diversas áreas [Breiman 2001, Cortes and Vapnik 1995, Cox 1958, Rumelhart et al. 1986], o que justifica sua adoção no presente trabalho.

Com o avanço da Inteligência Artificial, abordagens baseadas em aprendizado profundo têm ampliado as possibilidades de análise de imagens de sensoriamento remoto agrícola. [da Silva et al. 2024] apresentam um panorama dessas técnicas em cinco aplicações principais, destacando que arquiteturas convolucionais combinadas com séries temporais de imagens de satélite apresentam melhor desempenho do que abordagens que ignoram a estrutura espacial e temporal dos dados. O presente trabalho se posiciona como um estudo que opera sem essas informações, evidenciando seus limites e motivando sua incorporação como trabalho futuro.

Diferentemente dos estudos anteriores, este trabalho realiza uma comparação controlada entre um baseline de limiar de NDVI e algoritmos de ML supervisionados, utilizando conjuntos de desenvolvimento e teste extraídos de talhões distintos, com foco em discutir por que abordagens simples podem ser competitivas.

## 3. Conceitos e Técnicas

Esta seção apresenta os principais conceitos e técnicas utilizados neste trabalho.

### 3.1. Índice NDVI e Bandas Espectrais do Sentinel-2

O Índice de Vegetação por Diferença Normalizada (NDVI) é amplamente utilizado no monitoramento de culturas agrícolas a partir de imagens de satélite [Jensen 2009], sendo calculado conforme a Equação 1 [Galvão 2024] e adaptada para as bandas capturadas pelo Sentinel-2:

$$NDVI = \frac{B08 - B04}{B08 - B04} \tag{1}$$

em que B08 é a refletância do infravermelho próximo (842 nm) e B04 é a refletância do vermelho visível (665 nm). Valores próximos de 1 indicam vegetação densa e saudável; valores próximos de $-1$ indicam ausência de vegetação [Jensen 2009, ESA 2015, Florenzano 2011]. Neste trabalho, foram extraídas cinco estatísticas do NDVI por célula de 30m $\times$ 30m: média (*ndvi_mean*), desvio padrão (*ndvi_stdev*), mínimo (*ndvi_min*), máximo (*ndvi_max*) e amplitude (*ndvi_range*).

### 3.2. Limiar de Classificação (*Threshold*)

Um limiar $\tau$ aplicado ao *ndvi_mean* define uma regra simples: células com valor abaixo de $\tau$ são classificadas como falha; as demais, como normal [Huang et al. 2019]. O valor ótimo de $\tau$ foi determinado sobre o conjunto de desenvolvimento, maximizando o F1-score da classe falha, e aplicado diretamente ao conjunto de teste sem reajuste.

### 3.3. Otimização de Hiperparâmetros por Grid Search

O Grid Search testa sistematicamente todas as combinações possíveis de valores candidatos para cada hiperparâmetro, selecionando a configuração com melhor desempenho em validação cruzada [Ham et al. 2025]. Apesar do alto custo computacional, o método é determinístico e garante reprodutibilidade. Neste trabalho, foi aplicado com validação cruzada estratificada de 5 partições (*StratifiedKFold*, $k = 5$) sobre o conjunto de desenvolvimento.

### 3.4. Modelos de Aprendizado de Máquina

Os quatro algoritmos utilizados estão entre os mais empregados em problemas de classificação com dados de sensoriamento remoto [Reis et al. 2018]. A Tabela 1 apresenta uma síntese de cada um.

**Tabela 1. Modelos de Aprendizado de Máquina utilizados nos experimentos.**

| Algoritmo | Descrição |
|---|---|
| Random Forest | Conjunto de árvores de decisão treinadas sobre subamostras dos dados; a predição final é obtida por votação entre as árvores [Breiman 2001]. |
| SVM | Busca o hiperplano de maior margem entre as classes, podendo utilizar funções de kernel para dados não linearmente separáveis [Cortes and Vapnik 1995]. |
| MLP | Rede neural com camadas de entrada, ocultas e saída, treinada por retropropagação do erro [Rumelhart et al. 1986]. |
| Regressão Logística | Estima a probabilidade de pertencimento a uma classe por meio de uma função logística aplicada a uma combinação linear dos atributos [Cox 1958]. |

Modelos sensíveis à escala (SVM, MLP e Regressão Logística) foram aplicados com padronização prévia (*StandardScaler*). O Random Forest foi utilizado sem normalização. O desbalanceamento entre as classes foi tratado pelo parâmetro *class_weight="balanced"* em todos os modelos.

## 4. Metodologia

Esta seção descreve as etapas realizadas neste trabalho, desde a aquisição das imagens até a avaliação dos modelos. A Figura 1 ilustra o fluxo geral de desenvolvimento.

### 4.1. Aquisição de Dados e Área de Estudo

A área de estudo localiza-se no estado de Mato Grosso do Sul, em uma propriedade destinada ao cultivo comercial de eucalipto. A escolha do eucalipto como cultura-alvo é motivada pela relevância econômica da espécie na região: o Vale da Celulose de Mato Grosso do Sul concentra grandes plantações de multinacionais do setor de papel e celulose, que constituem uma das principais atividades econômicas do estado.

Foram selecionados dois talhões com variabilidade na densidade da cobertura vegetal, condição adequada para gerar amostras das classes normal e falha. As imagens foram obtidas na plataforma Copernicus Hub, utilizando o satélite Sentinel-2, com aquisição em abril de 2024. Ambos os talhões foram imageados na mesma data, garantindo condições equivalentes entre os conjuntos de desenvolvimento e teste. A área destacada em vermelho na Figura 2 indica a localização aproximada dos talhões selecionados para o estudo.

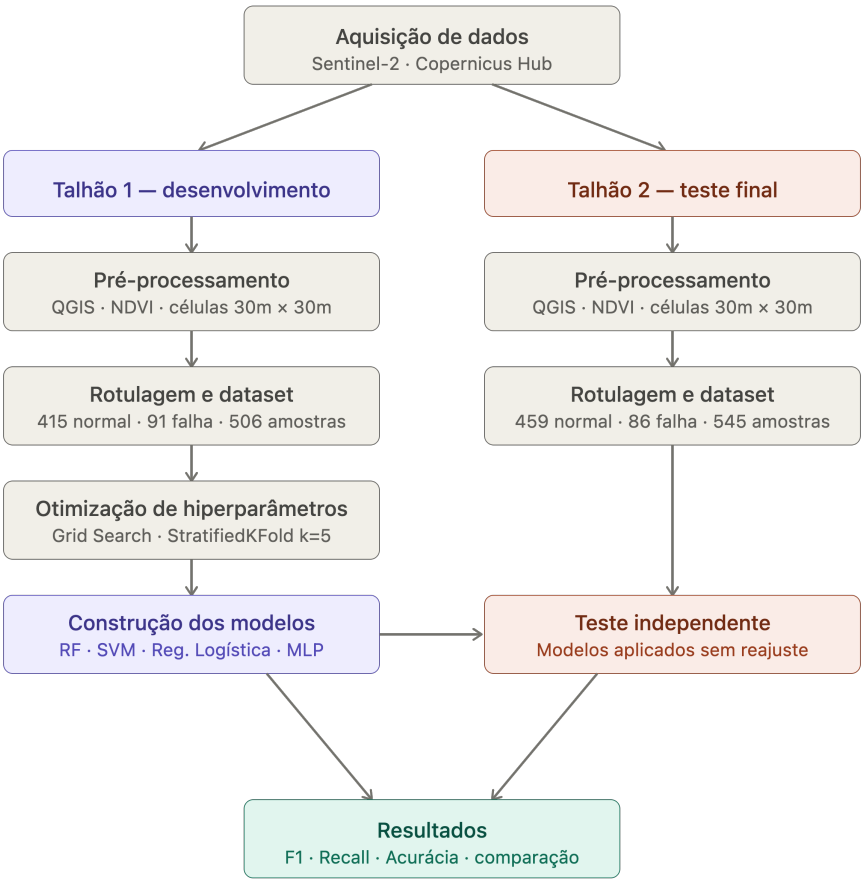



**Figura 1. Fluxo de desenvolvimento do projeto**



**Figura 2. Imagem de satélite da área de estudo escolhida**

### 4.2. Pré-Processamento de Imagens e Cálculo do NDVI

O pré-processamento das imagens foi realizado no software QGIS (versão 3.44.8). Os dois talhões foram delimitados manualmente por meio de polígonos vetoriais. Em seguida, cada talhão foi segmentado em células regulares de 30m $\times$ 30m, resolução compatível com a escala de análise do problema e com a resolução espacial das bandas Sentinel-2 utilizadas (10m/pixel), permitindo que cada célula contenha múltiplos pixels para o cálculo de estatísticas representativas da cobertura vegetal local.

Após a segmentação, o NDVI foi calculado pixel a pixel sobre cada talhão a partir das bandas B08 (842 nm) e B04 (665 nm) do Sentinel-2, conforme a Equação 1. O resultado desta etapa é apresentado na Figura 3.

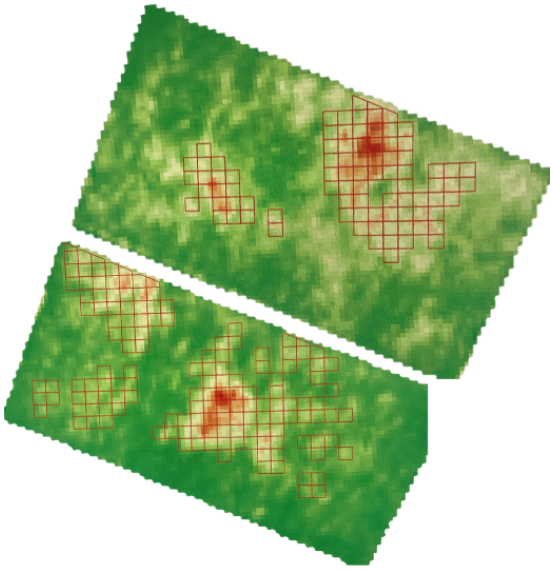


**Figura 3. Imagem dos talhões após aplicação do NDVI e Grid**

### 4.3. Construção do Dataset e Rotulagem

Para cada célula de 30m $\times$ 30m, foram extraídas estatísticas descritivas das bandas espectrais e do NDVI calculado. As células foram então rotuladas manualmente no QGIS como *normal* (classe 0) ou *falha* (classe 1), com base em inspeção visual das imagens de satélite e do mapa de NDVI. Os critérios visuais adotados para identificar regiões com falha foram baseados na experiência metodológica de [Galvão 2024], que identificou e categorizou falhas em talhões de eucalipto por meio de análise visual de imagens de satélite comparada a registros fotográficos de campo.

Os atributos extraídos por célula foram organizados em quatro grupos, conforme a Tabela 2.

Os dois talhões foram mantidos como conjuntos estritamente independentes: o primeiro foi utilizado para desenvolvimento dos modelos e o segundo foi reservado para avaliação final. A distribuição de amostras é apresentada na Tabela 3.

Os dados apresentam desbalanceamento entre as classes, com proporção aproximada de 4,6:1 e 5,3:1 entre normal e falha nos conjuntos de desenvolvimento e teste, respectivamente. Esse desbalanceamento reflete a condição real dos talhões e foi tratado

**Tabela 2. Grupos de atributos utilizados nos experimentos.**

| Grupo | Atributos |
|---|---|
| Grupo 1 | Estatísticas derivadas do NDVI |
| Grupo 2 | Média da banda B04 |
| Grupo 3 | Média da banda B08 |
| Grupo 4 | NDVI + B04 + B08 |

**Tabela 3. Distribuição de amostras por conjunto e classe.**

| Conjunto | Normal | Falha | Total |
|---|---|---|---|
| Desenvolvimento (Talhão 1) | 415 | 91 | 506 |
| Teste final (Talhão 2) | 459 | 86 | 545 |

em todos os modelos pelo parâmetro *class_weight="balanced"* da biblioteca scikit-learn (versão 1.7.2).

### 4.4. Baseline NDVI

Como referência comparativa, adotou-se um modelo de limiar baseado no atributo *ndvi_mean* de cada célula, adaptado da metodologia de [Huang et al. 2019]. O critério de classificação consiste em prever *falha* para toda célula com NDVI médio inferior a um limiar crítico $\tau$, e *normal* caso contrário.

O limiar ótimo foi determinado de forma empírica por busca exaustiva sobre 300 valores no intervalo $[-1{,}0;\ 1{,}0]$, maximizando o F1-score da classe falha no conjunto de desenvolvimento. O valor obtido foi $\tau = 0{,}7391$. Uma vez definido, esse limiar foi aplicado diretamente ao conjunto de teste independente, sem reajuste. A Figura 4 apresenta a curva de F1-score em função do limiar.

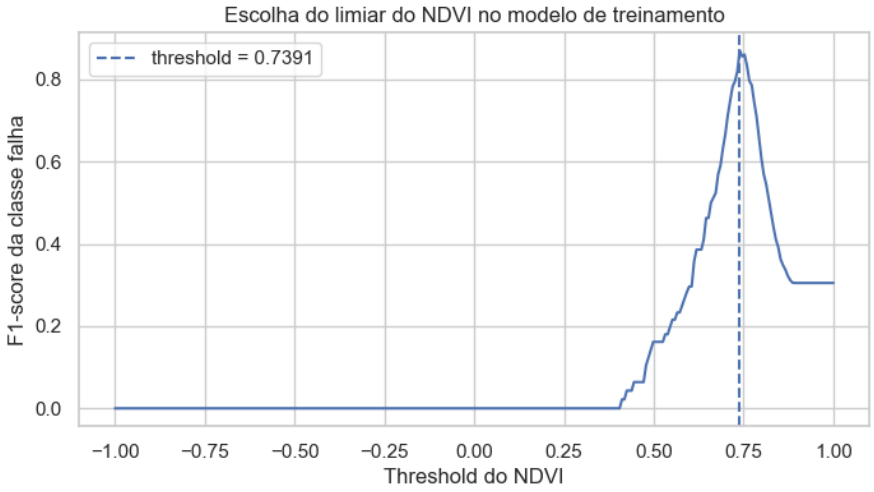



**Figura 4. Seleção do limiar ótimo de NDVI no conjunto de desenvolvimento.**

### 4.5. Modelos de Aprendizado de Máquina e Otimização de Hiperparâmetros

Quatro algoritmos supervisionados foram avaliados: Random Forest (RF), Support Vector Machine (SVM), Regressão Logística e Multi-Layer Perceptron (MLP). Para cada algo-

ritmo, modelos sensíveis à escala (SVM, MLP e Regressão Logística) foram aplicados com padronização prévia dos atributos (*StandardScaler*). O Random Forest foi utilizado sem normalização.

Os hiperparâmetros foram otimizados por Grid Search com validação cruzada estratificada de 5 partições (*StratifiedKFold*, $k = 5$), utilizando o F1-score da classe falha como critério de seleção, integralmente sobre o conjunto de desenvolvimento. O conjunto de teste foi mantido isolado durante toda essa etapa. Ao total, foram geradas e avaliadas 16 configurações (4 algoritmos $\times$ 4 grupos de atributos).

## 5. Resultados

A análise dos resultados foi organizada em torno de cinco perguntas formuladas a partir da hipótese do trabalho:

1. O Baseline NDVI foi suficiente ou foi superado pelos modelos de ML?
2. Qual algoritmo de ML apresentou o melhor F1-score para a classe falha?
3. Quais grupos de atributos produziram o melhor desempenho?
4. As bandas B04 e B08 trouxeram ganho real em relação ao NDVI isolado?
5. A troca de imagem entre desenvolvimento e teste afetou a generalização dos modelos?

A Tabela 4 apresenta as cinco melhores configurações avaliadas no conjunto de teste independente, ordenadas pelo F1-score da classe falha.

**Tabela 4. Top 5 configurações de modelos selecionadas por desempenho no teste final.**

| Pos. | Algoritmo | Grupo de Atributos | F1 | Recall | Acurácia | Precision (%) |
|---|---|---|---|---|---|---|
| 1 | Baseline NDVI | Grupo 1 – NDVI | 0,7870 | 0,9884 | 0,9156 | 0,6538 |
| 2 | MLP | Grupo 2 – B04 | 0,7788 | 0,9419 | 0,9156 | 0,6639 |
| 3 | MLP | Grupo 1 – NDVI | 0,7692 | 0,9884 | 0,9064 | 0,6296 |
| 4 | MLP | Grupo 4 – NDVI+B04+B08 | 0,7679 | 1,0000 | 0,9046 | 0,6231 |
| 5 | Random Forest | Grupo 4 – NDVI+B04+B08 | 0,7424 | 0,9884 | 0,8917 | 0,5944 |

### 5.1. O Baseline NDVI foi suficiente ou foi superado?

O Baseline NDVI obteve o maior F1-score entre todas as configurações avaliadas (0,7870), superando todos os modelos de ML no conjunto de teste independente. Esse resultado responde diretamente à hipótese do trabalho: quando os modelos utilizam apenas estatísticas simples do NDVI como entrada, algoritmos complexos não apresentam vantagem sobre uma regra de limiar. Isso ocorre porque o NDVI já concentra o principal sinal para distinguir regiões saudáveis de regiões com falha, e as estatísticas extraídas por célula não acrescentam informações suficientes para que os modelos de ML encontrem padrões mais complexos do que um simples corte de valor.

### 5.2. Qual algoritmo de ML apresentou o melhor desempenho?

Entre os modelos supervisionados, o MLP obteve os melhores resultados, ocupando as três primeiras posições no ranking de ML. O melhor resultado foi obtido com os atributos do Grupo 2 (apenas *b04_mean*), atingindo F1 = 0,7788 e acurácia de 91,56%. O Random Forest com o Grupo 4 ocupou a quinta posição geral (F1 = 0,7424). SVM e Regressão Logística apresentaram desempenhos inferiores em todas as configurações avaliadas.

### 5.3. Quais grupos de atributos produziram o melhor desempenho?

Os atributos do Grupo 1 (NDVI) e do Grupo 2 (B04) produziram os melhores resultados individuais. O Grupo 3 (apenas *b08_mean*) foi o de pior desempenho em todos os algoritmos, com F1-scores inferiores a 0,65, indicando que a banda B08 isolada tem menor capacidade de separar as classes neste problema. O Grupo 4, que combina todos os atributos, produziu resultados intermediários, sugerindo que a adição de atributos menos informativos pode reduzir o desempenho em alguns algoritmos.

### 5.4. As bandas B04 e B08 trouxeram ganho real?

A inclusão das bandas B04 e B08 no Grupo 4 não produziu melhora expressiva no F1-score em relação ao NDVI isolado. No entanto, o MLP com o Grupo 4 atingiu recall perfeito (1,0000), ou seja, nenhuma falha deixou de ser detectada. Em aplicações práticas de detecção de falhas agrícolas, não detectar uma falha real tende a ser mais prejudicial do que gerar um alarme falso. Por isso, dependendo do contexto de uso, o Grupo 4 pode ser preferível mesmo sem ganho em F1.

A análise de importância de atributos do Random Forest, apresentada na Figura 5, confirmou que os atributos do NDVI — especialmente *ndvi_mean* e *ndvi_min* — são os mais relevantes para as decisões do modelo, enquanto B04 e B08 aparecem com importância secundária.

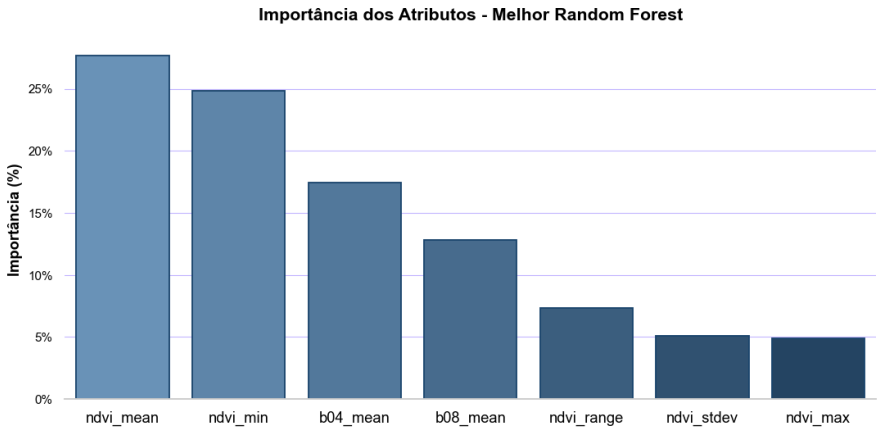



**Figura 5. Importância dos atributos estimada pelo Random Forest (Grupo 4 – NDVI+B04+B08).**

### 5.5. A troca de imagem afetou a generalização dos modelos?

As métricas obtidas no conjunto de teste independente (Talhão 2) foram consistentes com o desempenho de validação cruzada no conjunto de desenvolvimento (Talhão 1), sem quedas abruptas de desempenho. Isso indica que os modelos conseguiram generalizar para um talhão diferente do utilizado no treinamento, o que fortalece a validade dos resultados obtidos.

## 6. Conclusão

Este trabalho avaliou o uso de imagens do satélite Sentinel-2 e algoritmos de Aprendizado de Máquina para detectar falhas de plantio em talhões de eucalipto no estado de Mato

Grosso do Sul. Foram comparados quatro algoritmos — Random Forest, SVM, Regressão Logística e Multi-Layer Perceptron — com um baseline de limiar dinâmico de NDVI, utilizando atributos espectrais organizados em quatro grupos e avaliados em conjuntos de desenvolvimento e teste provenientes de talhões distintos.

O principal resultado foi que o Baseline NDVI superou todos os modelos de ML, alcançando F1-score de 0,7870 no conjunto de teste independente. Esse resultado responde diretamente à hipótese investigada: quando os modelos utilizam apenas estatísticas simples do NDVI como entrada, algoritmos mais complexos não apresentam vantagem sobre uma regra de limiar básica. Isso ocorre porque o NDVI já concentra o principal sinal para distinguir regiões saudáveis de regiões com falha, e os atributos utilizados não fornecem informação suficiente para que os modelos de ML encontrem padrões mais complexos.

Entre os algoritmos de ML, o MLP obteve o melhor desempenho isolado (F1 = 0,7788, Grupo 2 — banda B04). A inclusão das bandas B04 e B08 não gerou ganho expressivo em F1, mas produziu recall perfeito em algumas configurações, o que pode ser relevante em aplicações onde não detectar uma falha real é mais prejudicial do que gerar um alarme falso. A análise de importância de atributos do Random Forest confirmou o NDVI médio e o NDVI mínimo como as variáveis mais determinantes, reforçando que o sinal discriminativo do problema está concentrado no NDVI. A consistência das métricas entre os conjuntos de desenvolvimento e teste indicou boa capacidade de generalização dos modelos para dados de um talhão diferente do utilizado no treinamento.

As principais limitações do trabalho estão na forma como os dados foram representados: cada célula foi tratada de forma independente, sem considerar o estado das células vizinhas nem a variação temporal do NDVI ao longo do ciclo de crescimento do eucalipto. É exatamente essa ausência de informação espacial e temporal que mantém o problema simples o suficiente para que um limiar básico seja competitivo. Além disso, a rotulação baseada em NDVI introduz um possível viés e incerteza espacial, podendo conferir uma vantagem a favor da linha base do índice espectral, embora a identificação visual seja fundamentada em referências consolidadas.

Como trabalho futuro, propõe-se incorporar duas dimensões complementares. A primeira é o contexto espacial: incluir estatísticas das células vizinhas como atributos adicionais, ou utilizar Redes Neurais Convolucionais (CNNs) capazes de processar diretamente os mapas de NDVI como imagens. A segunda é a análise multitemporal: incorporar séries temporais de NDVI ao longo do ciclo vegetativo, permitindo capturar não apenas o estado pontual da vegetação, mas sua trajetória de desenvolvimento. A combinação dessas duas abordagens tem potencial para enriquecer o espaço de atributos e explorar melhor as capacidades dos algoritmos de ML [da Silva et al. 2024].

## Referências


Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):215–232.

da Silva, M. P., Schaefer, M. A., Oliveira, H. N., Reis, J. C., Nunes, I. M., and dos Santos, J. A. (2024). Learning visual patterns in remote sensing: An overview of agricultural applications. In *2024 37th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 1–6. IEEE.

ESA (2015). Sentinel-2 user handbook. Issue 1, revision 2, European Space Agency.

Florenzano, T. G. (2011). *Iniciação em Sensoriamento Remoto*. Oficina de Textos, São Paulo, 3 edition.

Galvão, M. d. F. (2024). Sensoriamento remoto para identificação de falhas em cultura do eucalipto. Trabalho de Conclusão de Curso (Graduação) - Engenharia Agrícola e Ambiental, Universidade Federal de Viçosa, Viçosa.

Ham, A. G., Nafornita, C., Vesa, V. C., Copacean, G., Davidovici, V. D., and Nafornita, I. (2025). Grid search and genetic algorithm optimization of neural networks for automotive radar object classification. *Sensors*, 25(19):6017.

Huang, X., Liu, J., Zhu, W., Atzberger, C., and Liu, Q. (2019). The optimal threshold and vegetation index time series for retrieving crop phenology based on a modified dynamic threshold method. *Remote Sensing*, 11(23):2725.

Jensen, J. R. (2009). *Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres*. Parêntese Editora, São José dos Campos, SP. Tradução de José Carlos Neves Epiphanio.

Leivas, J. F., Teixeira, A. H. d. C., Takemura, C. M., and Garçon, E. A. M. (2022). Análise da dinâmica temporal de índices de vegetação NDVI, EVI, SAVI e IRECI através de imagens Sentinel-2a e MODIS. In *Ciências Agrárias: o avanço da ciência no Brasil*, volume 4, pages 334–343. Editora Científica Digital.

Massruhá, S. M. F. S., LEITE, M. A. d. A., OLIVEIRA, S. R. d. M., MEIRA, C. A. A., LUCHIARI JUNIOR, A., and BOLFE, É. L., editors (2020). *Agricultura digital: pesquisa, desenvolvimento e inovação nas cadeias produtivas*. Embrapa, Brasília, DF.

Reis, A. A. d., Carvalho, M. C., de Mello, J. M., Gomide, L. R., Ferraz Filho, A. C., and Acerbi Junior, F. W. (2018). Spatial prediction of basal area and volume in Eucalyptus stands using Landsat TM data: an assessment of prediction methods. *New Zealand Journal of Forestry Science*, 48(1):1–17.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.

Santos, E. K. M. d., Figueiredo, S. M. d. M., and Figueiredo, E. O. (2020). Comportamento fenológico de espécies florestais detectado por sistemas aéreos não tripulados. In *Biodiversidade e Biotecnologia no Brasil 1*, chapter 17, pages 260–277. Editora Stricto Sensu.