

USO DE MACHINE LEARNIG PARA CLASSIFICAÇÃO DE ESTRESSE HÍDRICO NO TRIGO

**JOSÉ WILSON DE OLIVEIRA MAGALHÃES¹, BIANCA BATISTA BARRETO²,
BIANCA BATISTA BARRETO⁴, JOÃO LEONARDO FERNANDES PIRES³, LUÍS
HENRIQUE BASSOI⁵**

¹ Eng. Agrônomo, Mestrando em Engenharia Agrícola, FCA/Unesp, Botucatu – SP, jw.magalhaes@unesp.br.

² Graduanda em Biotecnologia, bolsista de iniciação científica do CNPq, UFScar, São Carlos – SP.

³ Eng. Agrônomo, Pesquisador, Embrapa Trigo, Passo Fundo – RS.

⁴ Eng. Agrônômica, Pesquisadora, Embrapa Instrumentação, São Carlos – SP

⁵ Eng. Agrônomo, Pesquisador, Embrapa Instrumentação, São Carlos – SP.

Apresentado no

LII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2023
18 a 21 de outubro de 2023 – Ribeirão Preto - SP, Brasil

RESUMO: O uso de algoritmos de Machine Learning (ML) proporciona uma robusta capacidade no processamento e análise de dados. As técnicas de ML fornecem respostas em menor tempo comparadas aos modelos de regressão estatística. O experimento com o cultivo do trigo foi conduzido no período de 13 de outubro de 2022 a 16 de janeiro de 2023 em casa de vegetação do Laboratório de Referência Nacional em Agricultura de Precisão (LANAPRE), Embrapa Instrumentação. O delineamento utilizado foi o inteiramente casualizado, com 16 repetições, tendo como fatores o manejo de irrigação (irrigação plena e irrigação com déficit hídrico) e foram semeadas duas cultivares de trigo (BRS 394 e BRS 264). A temperatura média das plantas apresentou uma diferença entre os manejos de irrigação de aproximadamente 1°C, distinguindo, com o auxílio das imagens térmicas, plantas em estresse hídrico. O algoritmo de classificação mostrou-se uma ótima ferramenta para classificação e identificação do estresse hídrico nas plantas de trigo com uma taxa de acerto de 90%.

PALAVRAS-CHAVE: *Triticum spp*, algoritmo, imagem térmica.

USE OF MACHINE LEARNIG FOR CLASSIFICATION OF WATER STRESS IN WHEAT

ABSTRACT: The use of Machine Learning (ML) algorithms provides a robust capability in data processing and analysis. ML techniques provide answers in less time compared to statistical regression models. The experiment with wheat cultivation was conducted from October 13, 2022 to January 16, 2023 in a greenhouse at the National Reference Laboratory in Precision Agriculture (LANAPRE), Embrapa Instrumentation. The design used was completely randomized, with 16 replications, with irrigation management as factors (full irrigation and irrigation with water deficit) and two wheat cultivars (BRS 394 and BRS 264) were sown. The average temperature of the plants presented a difference between the irrigation managements of approximately 1°C, distinguishing, with the aid of thermal images, plants under water stress. The classification algorithm proved to be a great tool for classifying and identifying water stress in wheat plants with a success rate of 90%.

KEYWORDS: *Triticum spp*, algorithms, thermal images..

INTRODUÇÃO: O termo Machine Learning vem no inglês e pode ser entendido como a capacidade dos computadores de aprenderem com base em dados disponíveis. Os algoritmos são classificados em algoritmos de classificação e regressão. A classificação gera resultados categóricos (ou discretos), enquanto a regressão resulta em valores contínuos. Os algoritmos de Machine Learning também podem ser classificados em aprendizado supervisionado e não supervisionado. No primeiro, para cada amostra de entrada é também apresentado ao algoritmo o resultado esperado na saída, conhecido também como rótulo. Existem diversos algoritmos de aprendizado supervisionado, entre eles: K-Nearest Neighbors (K-NN), Regressão Linear Polinomial e Logística, Support Vector Machine (SVM), Decision Tree, Random Forest (RF), Light GBM, XGBoost e Redes Neurais Artificiais (RNA). No aprendizado não supervisionado, apenas as amostras de entrada são apresentadas ao algoritmo, e os dados utilizados para ajustar (treinar) os modelos rotulados não são rotulados. Dessa forma, espera-se que o algoritmo classifique os dados em grupos. Entre os algoritmos, podem-se citar os algoritmos de cluster (agrupamento), como: K-Means, Fuzzy K-Means e Agrupamento Hierárquico (aglomerativo, divisivo). Modelos de Machine Learning podem utilizar dados categóricos e numéricos (binários, inteiros, reais) provenientes de diferentes tipos de sensores. Esses dados são empregados para treinar (ajustar) o modelo. O banco de dados é composto de observações (amostras) e atributos (características). As observações são organizadas em linhas no banco de dados e os atributos, em colunas. Esse formato é padrão utilizado nos algoritmos de aprendizagem de máquinas (MENEZES, 2021). Assim, o objetivo deste trabalho é criar modelos de aprendizado de máquina para identificar e classificar plantas de trigo submetidas a estresse hídrico em ambiente controlado sob irrigação localizada.

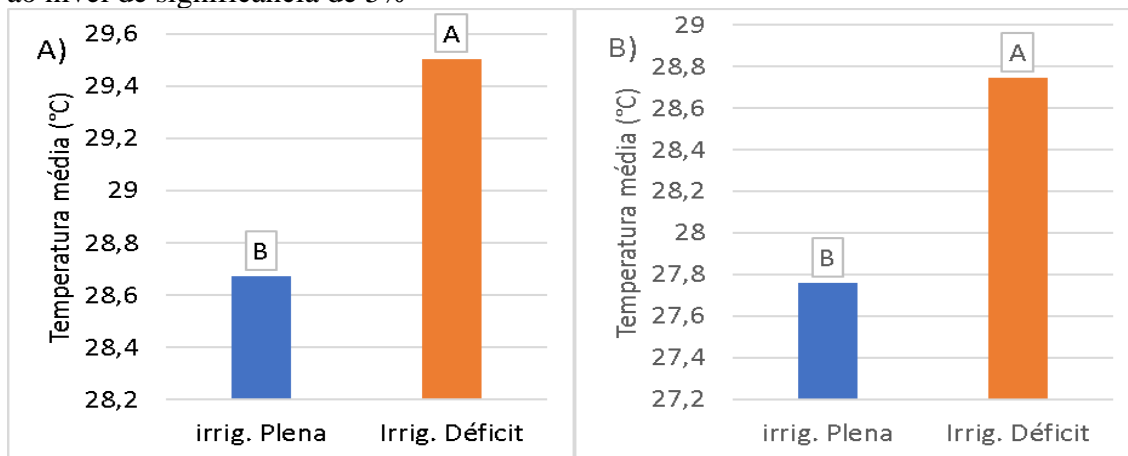
MATERIAL E MÉTODOS: O experimento foi desenvolvido na casa de vegetação do Laboratório de Referência Nacional em Agricultura de Precisão (LANAPRE), pertencente a Embrapa Instrumentação, em São Carlos – SP. Vasos com volume unitário de 15 dm³ foram utilizados para o cultivo de trigo em casa de vegetação. Os mesmos foram preenchidos com 1,5 kg de brita, uma manta geotêxtil para evitar a perda de solo pela drenagem e, por fim, 12 kg de terra fina seca ao ar (TFSA) passados em peneira de 2 mm até o solo atingir 0,02 m da borda, compactando-se de modo a manter sua densidade o mais uniforme possível. Drenos na lateral inferior do vaso permitiram a saída do excesso de água do solo. O solo utilizado no preenchimento dos vasos é classificado como Latossolo Vermelho Amarelo Distrófico. A granulometria está distribuída em 32% de areia, 6% de silte e 62% argila, o que conferiu uma textura franco-argilosa. O sistema de irrigação utilizado na casa de vegetação foi o gotejamento, com vazão de 1,6 L h⁻¹ por gotejador, com 4 emissores por vaso. Para o experimento, foram utilizadas duas cultivares de trigo, a BRS 394 e BRS 264, avaliadas separadamente. Em ambas as cultivares, dois manejos de irrigação foram realizados. A irrigação plena ocorreu durante todo o ciclo da cultura, com 4 emissores por vaso. A irrigação com déficit foi realizada com a interrupção da irrigação no estágio de aparecimento da folha bandeira e na maturação. O início do manejo de irrigação com déficit ocorreu na aparição da folha bandeira, aos 33 e 31 DAE (BRS 394 e BRS 264, respectivamente), com a interrupção da irrigação. A irrigação foi novamente realizada na fase de enchimento de grãos (41 e 44 DAE) e interrompida novamente aos 55 e 57 DAE, respectivamente para BRS 394 e BRS 264, até a colheita. A medição da termografia infravermelho foi realizada em casa de vegetação contendo os dois tratamentos na mesma imagem. A câmera utilizada foi da marca FLIR e modelo T-420 fixada em um tripé, de maneira que a distância até o chão fosse equivalente a 1,5 m. As medições foram iniciadas às 8:30 da manhã. Para a padronização da medição, os vasos foram posicionados sempre no mesmo local. Para a realização das medições, foram inseridos dados de entrada como a altura da câmera até o chão equivalente a 1,5 m; temperatura ambiente; umidade relativa e emissividade da planta equivalente a 0,96.

Os dados de temperatura e umidade foram obtidos com auxílio de um aparelho termo-higrômetro da marca Prolab. As imagens térmicas foram processadas no programa Flir Thermal Studio, na versão livre. Cada imagem termográfica foi dividida em duas áreas de avaliação (irrigação plena ao lado de irrigação com déficit, ambos com a mesma dose de hidrogel). Os dados foram adicionados em uma planilha, subdivididos em temperatura máxima, temperatura média e temperatura mínima de cada área de cada tratamento. Para minimizar o efeito da variação do valor da temperatura diária, para todos os valores de temperatura foi retirado o gradiente da temperatura ambiente, como indicado pela equação abaixo.

$$T = ((\text{Temperatura da área} - \text{Temperatura ambiente}) / (\text{Temperatura ambiente})) \quad (1)$$

Onde T = Temperatura da planta (°C). Os dados obtidos foram submetidos à análise de variância e aos testes de comparação de médias utilizando o software estatístico AgroEstat®, obedecendo o delineamento experimental inteiramente casualizado. Os dados foram analisados no software Weka versão 3.8.

RESULTADOS E DISCUSSÃO: Mediante a análise de variância das cultivares de trigo a temperatura média das plantas de ambas as cultivares apresentaram respostas aos tratamentos de irrigação, ao nível de significância de 5%. As médias foram submetidas ao teste de Tukey ao nível de significância de 5%



Médias seguidas pela mesma letra não se diferem estatisticamente segundo teste de Tukey ao nível de significância de 5%.

FIGURA 1. Médias temperatura das plantas das cultivares BRS 394(A) e BRS 264 (B)

Em ambas as cultivares a diferença na temperatura do dossel das plantas foi em torno de 1°C. Resultado similares foram observados em estudo de Duan et al. (2018), no qual as plantas de trigo sob condições de estresse hídrico apresentaram uma redução na transpiração, ocasionada pelo fechamento dos estômatos e que provocou a diminuição da transpiração, causando o aumento da temperatura do dossel da planta. Para um aumento de 1°C na temperatura do dossel da planta, estima-se a diminuição da transpiração em 10% (Helyes et al., 2010). O processamento dos dados, realizado através do Programa WEKA, versão 3.8, seguiu as etapas de padronização e classificação. A padronização dos dados visou o tratamento destes para viabilizar a aplicação de algoritmos de classificação, no contexto de aprendizagem de máquina, identificando a métrica e a normalização utilizada. Enquanto a classificação buscou por meio de modelos identificar plantas das diferentes classes estudadas, determinando o nível de diferenciação entre os tratamentos. Entre os modelos estudados, o que resultou em maior acurácia nos conjuntos de dados em questão, considerando-se as diferentes técnicas avaliadas, foi o denominado Random Forest.

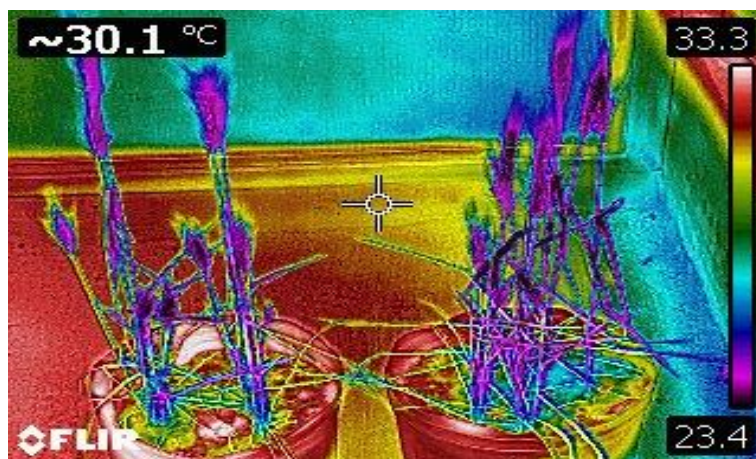


FIGURA 2. Imagem térmica capturada por meio de uma câmera termográfica

A tabela 1 apresenta os resultados as medições de temperatura da planta, incluindo taxa de acerto e erro e matriz de confusão com os valores de verdadeiro e falso positivo e negativo (classificação a e b). E o resumo das taxas de acerto para as técnicas, em que a termografia apresentou uma elevada taxa de acerto com o valor de 90%, para ambas as cultivares. Apesar das taxas de acertos para as cultivares terem sido relativamente altos, os seus respectivos valores apresentaram coerência entre si, sendo este resultado reflexo da indução do déficit hídrico durante o ciclo das plantas.

BRS 394	BRS 264
Correctly Classified Instances 90.000%	Correctly Classified Instances 90.000%
Incorrectly Classified Instances 10.000%	Incorrectly Classified Instances 10.000%
Confusion Matrix ===	Confusion Matrix ===
a b <-- classified as	a b <-- classified as
4 1 / a = IP	4 1 / a = IP
0 5 / b = ID	0 5 / b = ID

Tabela 1. Índices de acertos e erros em análise para as cultivares em que IP é irrigação Plena e ID é a irrigação com déficit

CONCLUSÕES: A classificação e identificação com algoritmo de classificação em plantas de trigo submetidas a déficit hídrico tem alta taxa de acerto com a captura de imagens térmicas por meio de câmera termográfica.

AGRADECIMENTOS: À CAPES pela concessão de bolsa de estudo ao primeiro autor; ao CNPq pelo apoio financeiro ao projeto; à Embrapa Instrumentação pelo apoio institucional e à Fertgel Hidrogéis e Fertilizantes Inteligentes pela parceria.

REFERÊNCIAS:

- DUAN, H.; CHASZAR, B.; LEWIS, J.D.; SMITH, R.A.; HUXMAN, T.E.; TISSUE, D.T. CO₂ and temperature effects on morphological and physiological traits affecting risk of drought-induced mortality. **Tree Physiology**, 38, 1138–1151, 2018.
- HELYES, L; BÓCS, A; PÉK, Z. Effect of water supply on canopy temperature, stomatal conductance and yield quantity of processing tomato (*Lycopersicon esculentum* Mill.). **International Journal Horticultural Science and Technology**, v. 16, 13–15. 2010
- MENESES, K. C. **Rumo à agricultura inteligente: previsão de produtividade agrícola com dados agrometeorológicos usando machine learning**. 2021. 97f. Tese (Doutorado em Agronomia, Área de concentração: Ciência do Solo) Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, São Paulo. 2021.