

Predição do início da estação chuvosa no Mato Grosso usando algoritmos de aprendizado de máquina

Matheus O. Lisboa; Daniel L. Fernandes; Marcos H. Costa; Lucas L. Fernandes; Gustavo S. Mairink

Dimensões Ambientais: ODS13

Introdução

O estado do Mato Grosso (MT), uma potência agrícola do Brasil, enfrenta um desafio crítico, a crescente imprevisibilidade do início da estação chuvosa (IEC), o que define diretamente o sucesso do plantio. Essa incerteza climática ameaça a sustentabilidade do agronegócio, impactando diretamente o planejamento da safra. Diante deste cenário, este projeto aplicou algoritmos de Aprendizado de Máquina (AM) para criar modelos precisos de previsão do IEC, buscando oferecer maior segurança e eficiência para a agricultura na região.

Objetivos

O objetivo geral deste trabalho foi desenvolver e avaliar um modelo de AM para prever, de forma espacialmente explícita, o IEC em cada região (*pixel*) do MT, classificando-o em intervalos de tempo. Além do objetivo principal, este trabalho propôs como objetivos específicos:

- Realizar a seleção das variáveis preditoras mais relevantes com base em testes de correlação linear;
- Estruturar e pré-processar o conjunto de dados para a modelagem;
- Aplicar diferentes algoritmos classificação multiclasse;
- Comparar o desempenho dos modelos por meio de métricas de avaliação e testes estatísticos, identificando a melhor performance;
- Analisar a distribuição espacial dos acertos e erros do modelo final.

Metodologia

- **Alvo da Predição:** O IEC (1981-2024) classificado em 5 intervalos (P1-P5), a partir de dados de precipitação do CHIRPS para 303 *pixels* por ano no MT com resolução de 0,5° (~55 km), totalizando em 13.332 alvos;
- **Variáveis Preditoras:** As coordenadas geográficas (Lat/Lon) e a anomalia da média mensal da Temperatura (em Kelvin) da Superfície do Mar (TSM) em três regiões do oceano Atlântico (R1, R2 e R3) para o mês de fevereiro, obtidas por meio da reanálise ERA5;
- **Modelagem:** Foram aplicados 6 algoritmos de AM supervisionado ao conjunto de dados, com as otimizações dos seus principais hiperparâmetros por meio da técnica *Bayes Search*;
- **Treinamento e Avaliação:** Os modelos criados foram treinados e avaliados por meio da validação cruzada estratificada em 10 partes, e o desempenho foi mensurado com base em 6 métricas de avaliação;
- **Validação Estatística:** Os resultados dos modelos foram comparados estatisticamente por meio dos testes de Shapiro-Wilk, ANOVA e de Tukey para a seleção do modelo final.

Apoio Financeiro

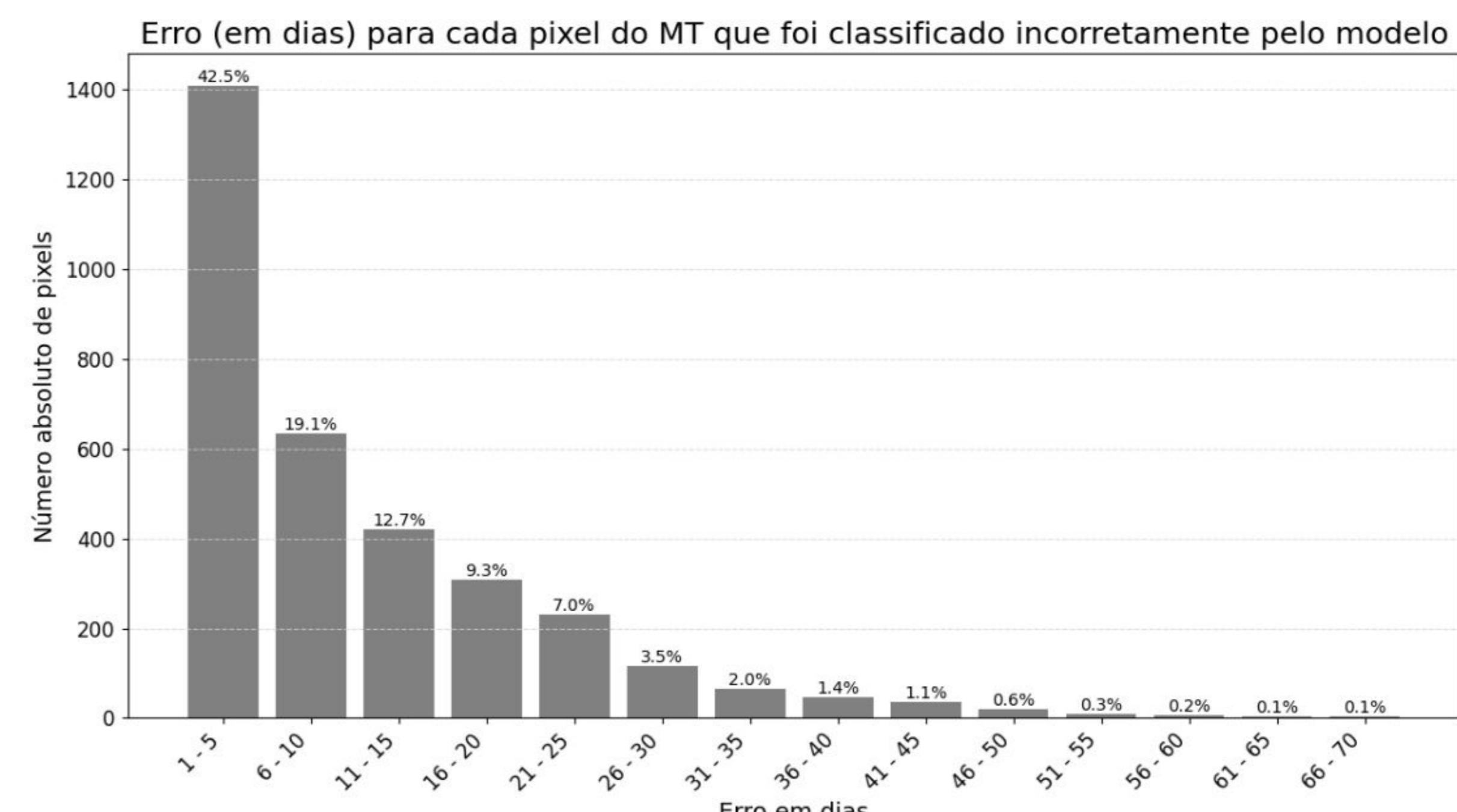
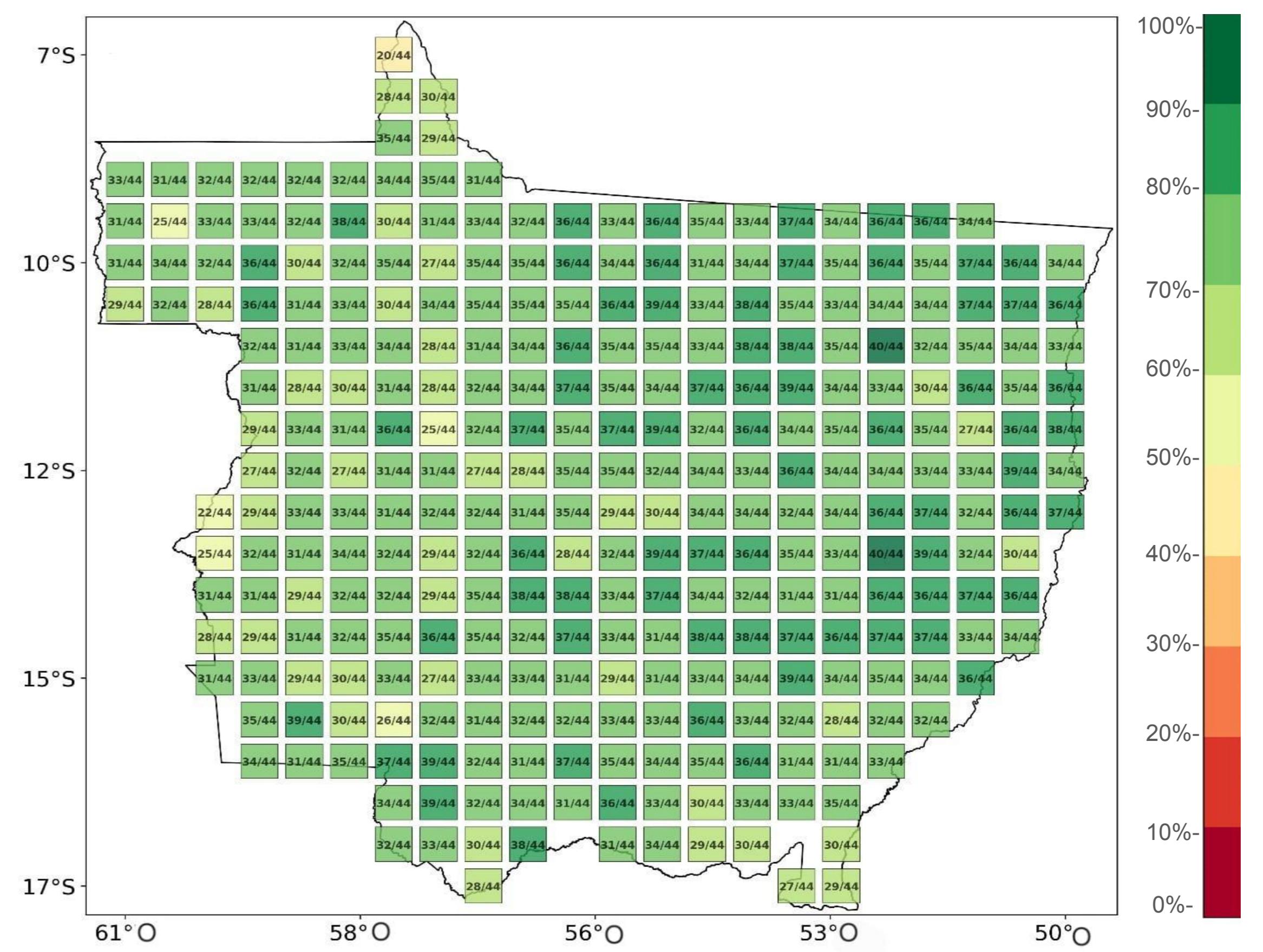
Resultados

Tabela 1: Métricas de Desempenho

Modelo	ACC (%)	F1 (%)	AUC	MCC	T. Ex. (s)
CatBoost	75,15	75,07	0,943	0,661	34,56 ± 0,29
Decision Tree	70,63	70,63	0,803	0,602	—
Gaussian Naive Bayes	43,07	40,55	0,728	0,201	—
LightGBM	74,51	74,45	0,940	0,653	—
Random Forest	75,02	74,94	0,941	0,659	13,20 ± 0,71
XGBoost	75,47	75,40	0,940	0,666	1,79 ± 0,74

Tabela 2: Matriz de Confusão do XGBoost

Classes verdadeiras	Classes preditas					Total de Pixels
	P1	P2	P3	P4	P5	
P1	829	220	96	31	12	1.188
P2	207	2.074	431	98	21	2.831
P3	92	387	4.438	359	125	5.401
P4	43	69	528	1.859	149	2.648
P5	8	31	176	187	962	1.364



Conclusões

A aplicação de algoritmos de AM, com destaque para o XGBoost, revelou-se uma abordagem promissora para a previsão do IEC no MT. Essa metodologia oferece suporte quantitativo relevante à tomada de decisão no agronegócio, contribuindo para o planejamento agrícola e a mitigação de riscos climáticos.

Bibliografia

BRUMATTI, Livia Maria; PIRES, Gabrielle Ferreira; SANTOS, Ana Beatriz. *Challenges to the adaptation of double cropping agricultural systems in Brazil under changes in climate and land cover*. *Atmosphere*, v. 11, n. 12, p. 1310, 2020.

COMMAR, Luiz Felipe Sant'Anna *et al.* *Mato Grosso's rainy season: past, present, and future trends justify immediate action*. *Environmental Research Letters*, v. 19, n. 11, p. 114065, 2024.