



BIC/UCS

Inteligência Artificial Explicável XAI

Autores: Lucas Schiochet, Carine G. Webber

INTRODUÇÃO / OBJETIVO

A inteligência artificial é a nova eletricidade. Ela irá transformar todos os setores da sociedade e impulsionar o progresso humano de uma forma sem precedentes [1]. Atualmente, podemos perceber que a cada dia que passa a IA faz mais parte do nosso dia-a-dia, em aplicativos de compras, bancos, etc. No entanto, muitas pessoas relatam uma falta de confiança nessas tecnologias e nas empresas que as gerenciam e produzem [2]. Dessa necessidade surge a XAI, um método que tem sido estudado para tornar as previsões da IA em algo mais compreensível para o ser humano.

Nesse contexto, este trabalho tem como objetivos:

- Analisar um caso de abandono universitário e usar os dados para treinar uma rede neural;
- Verificar por meio da explicabilidade da IA quais os atributos que a IA julgou mais importante para chegar no resultado.

MATERIAL E MÉTODOS

A condução foi feita em três estágios: coleta, anonimização e preparação de dados dos estudantes de um curso de Ciências Biológicas EAD, com atributos como forma de ingresso, tipo de cota e distância do polo; o desenvolvimento de uma rede neural multicamada (MLP), para classificação entre abandono, cancelamento, formação e vínculo; a aplicação das bibliotecas Python SHAP[3] e LIME[4] para interpretar os saídas da rede neural; por fim a análise desses resultados para cada biblioteca e a comparação do desempenho e dos resultado de cada biblioteca.

RESULTADOS OU RESULTADOS ESPERADOS

Os resultados indicam que atributos como "Nota Final", "Distância do Polo" e "Forma de Ingresso" foram os mais relevantes para prever abandono universitário. As análises com SHAP mostraram que estudantes com menores notas finais e maior distância do polo apresentaram maior probabilidade de abandono, como mostra a

RESULTADOS OU RESULTADOS ESPERADOS

figura 1.

Já o LIME permitiu interpretações locais, identificando, por exemplo, como pequenas variações em atributos individuais impactam a previsão.

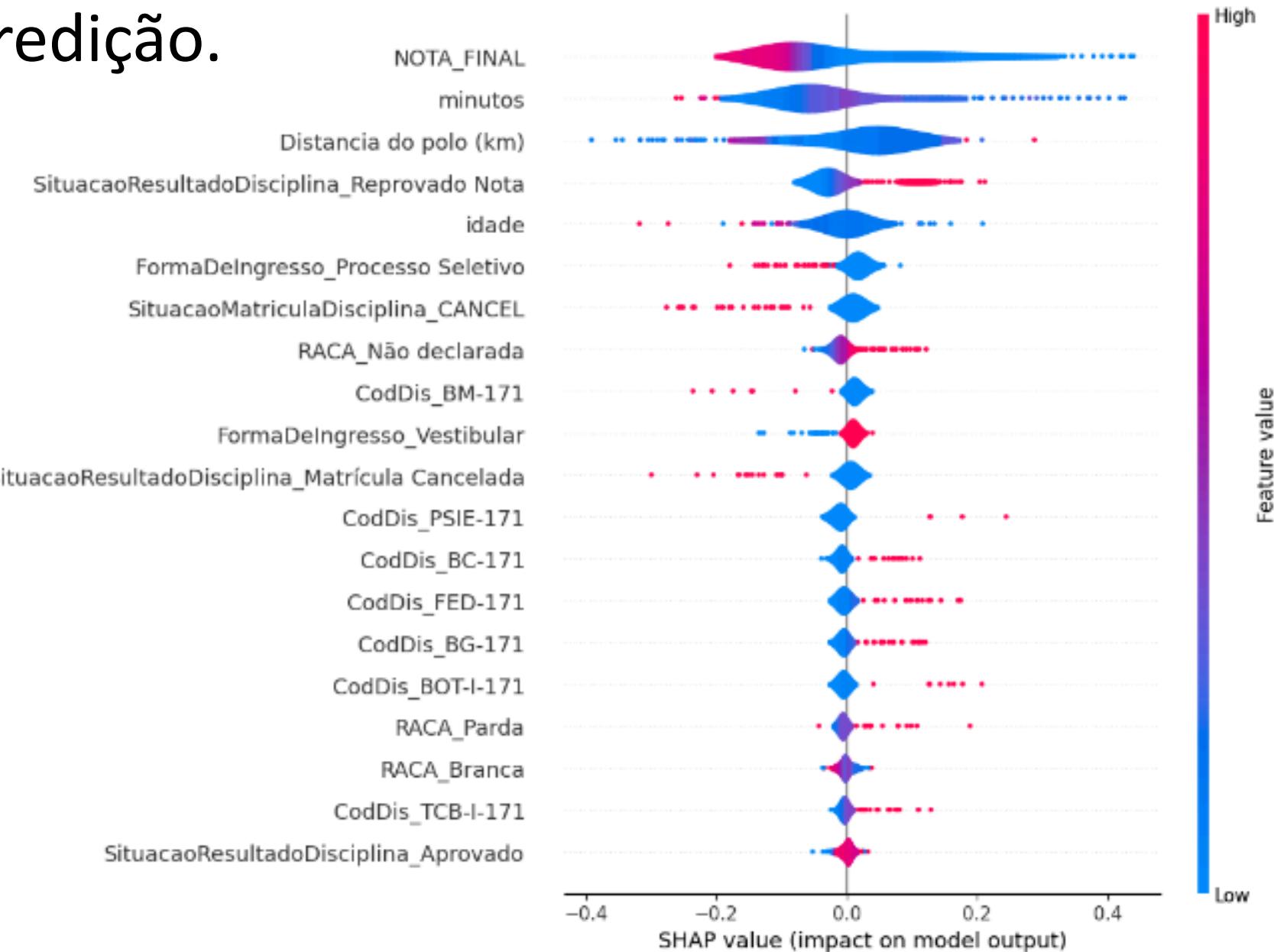


Figura 1. Gráfico SHAP tipo violino para abando

CONSIDERAÇÕES FINAIS

Uma limitação observada é o alto custo computacional do SHAP em comparação ao LIME. No entanto, a robustez das explicações geradas pelo SHAP justifica seu uso para compreensão global do modelo. Ambas as ferramentas demonstraram-se valiosas para a auditoria de sistemas de IA, aumentando a confiança e a transparência.

Também é importante ressaltar a importância da busca por transformar a IA cada vez mais transparentes, principalmente em áreas sensíveis.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] NG, Andrew. Artificial Intelligence is the New Electricity. [S.I.]: Stanford Graduate School of Business, 2017. Disponível em: <https://www.youtube.com/watch?v=21EiKfQYZXc>. Acesso em: 4 jun. 2025.
- [2] von Eschenbach, W. J. (2021). Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology*, 34, 1607-1622.
- [3] Disponível em: <https://shap.readthedocs.io/en/latest/>
- [4] Disponível em: <https://github.com/marcotcr/lime>