

Monitoramento de modelos na Mineração de Dados Educacionais: Impactos no desempenho e interpretabilidade conforme atualização dos dados

Otávio Santos Silva Corrêa¹, Cássio Soares Carvalho¹,
Marilton Sanchotene de Aguiar¹, Júlio Carlos Balzano de Mattos¹

¹Centro de Desenvolvimento Tecnológico – Universidade Federal de Pelotas (UFPEL)
Rua Gomes Carneiro, 1 – 96.010-610 – Pelotas – RS – Brasil

otavioscorrea@gmail.com, cassio.carvalho@inf.ufpel.edu.br

marilton@inf.ufpel.edu.br, julius@inf.ufpel.edu.br

Abstract. *Machine Learning (ML) techniques achieve high predictive performance; however, the “black-box” nature of these models limits their practical adoption. This study analyzes the interpretability and the impacts of temporal updates on ML models trained for predicting academic dropout. The models were trained using data from UFPEL over consecutive semesters, enabling temporal analysis. The results demonstrate that the continuous incorporation of data improves model performance, while the application of SHAP (SHapley Additive exPlanations) ensured explainability. In conclusion, combining evolutionary data and interpretability techniques promotes a more transparent and responsible use of Artificial Intelligence within the educational context.*

Resumo. *Técnicas de Aprendizado de Máquina (AM) alcançam alto desempenho preditivo, porém a natureza “caixa preta” desses modelos limita sua adoção prática. Este trabalho analisa a interpretabilidade e os impactos da atualização temporal em modelos de AM treinados para previsão de evasão acadêmica. Os modelos foram treinados utilizando dados da UFPEL ao longo dos semestres, possibilitando a análise temporal. Resultados mostram que incorporar dados continuamente melhora o desempenho dos modelos, enquanto a aplicação do SHAP (SHapley Additive exPlanations) garantiu a explicabilidade. Conclui-se que unir dados evolutivos e técnicas de interpretabilidade promove o uso mais transparente e responsável da Inteligência Artificial no contexto educacional.*

1. Introdução

A Mineração de Dados Educacionais (MDE) constitui um campo de investigação voltado ao desenvolvimento de métodos destinados à compreensão dos processos de aprendizagem e ao aprimoramento dos ambientes educacionais [Baker et al. 2010]. Trata-se de uma área interdisciplinar que, com o crescente desenvolvimento das inteligências artificiais, está em ascensão ao integrar técnicas como Aprendizado de Máquina (AM), estatística avançada, Mineração de dados, psicologia educacional e cognitiva [Xiao et al. 2022]. Um domínio de aplicação de alta relevância desse campo corresponde à predição de desempenho e à identificação precoce do risco de evasão dos alunos no ensino superior

[Xiao et al. 2022, Bakhshinategh et al. 2018], um problema de escala internacional que acarreta desperdícios sociais, acadêmicos e econômicos [Silva Filho et al. 2007].

Embora técnicas de Aprendizado de Máquina (AM) alcancem alto desempenho preditivo na identificação precoce de alunos em risco de evasão, a natureza “caixa preta” desses modelos limita sua adoção prática devido à falta de transparência [Carvalho et al. 2019, Linardatos et al. 2021] e à potencial introdução de vieses [Saxena et al. 2019]. Nesse cenário, surgem os conceitos de interpretabilidade e explicabilidade, frequentemente utilizados de forma intercambiável, que indicam o grau em que um ser humano, sem conhecimento absoluto do assunto, é capaz de compreender a causa de uma decisão ou prever, de maneira consistente, o resultado de um modelo [Molnar 2022].

Além da barreira da explicabilidade, sistemas de IA na educação enfrentam o desafio da dinamicidade inerente ao ambiente acadêmico. À medida que os semestres avançam, fatores institucionais, condições socioeconômicas e o próprio perfil comportamental dos estudantes sofrem alterações [Haug et al. 2022]. Na literatura, a mudança nas propriedades estatísticas das variáveis ao longo do tempo é denominada desvio de conceito (*concept drift*) [Gama et al. 2014, Lu et al. 2019]. Ignorar essa progressão pode fazer com que padrões aprendidos no passado percam a validade, levando o sistema a gerar diagnósticos de risco obsoletos ou incorretos [Neto et al. 2025], o que reforça a necessidade de monitoramento de modelos educacionais após implantação [Carvalho et al. 2026].

Este trabalho propõe investigar estratégias de monitoramento e atualização de modelos após a implantação, com o objetivo de avaliar como futuras modificações podem afetar a relação entre desempenho preditivo e interpretabilidade. No contexto da predição de evasão no ensino superior, portanto, busca-se responder:

- Q1 – A atualização contínua dos dados melhora o desempenho preditivo?
- Q2 – De que forma a atualização dos dados altera os fatores explicativos identificados pelo SHAP(SHapley Additive exPlanations)?
- Q3 – É possível detectar indícios de *concept drift* por meio das explicações?

O artigo está organizado da seguinte forma. A Seção 2 apresenta trabalhos envolvendo temas relacionados ou abordagens similares ao presente estudo. A Seção 3 descreve o *dataset* utilizado, a estratégia para simular a evolução temporal dos dados e o *pipeline* para a obtenção dos modelos de predição. A Seção 4 apresenta os resultados preliminares e, por fim, a Seção 5 apresenta as conclusões.

2. Trabalhos relacionados

Estudos para compreender e mitigar a evasão no ensino superior têm evoluído consideravelmente nas últimas décadas, transitando de análises estatísticas e demográficas de caráter macroscópico para abordagens preditivas refinadas, baseadas em MDE e AM.

No estudo pioneiro de [Silva Filho et al. 2007], mapeou-se, de forma macroscópica, as taxas anuais médias de abandono no país, evidenciando uma correlação negativa entre a demanda pelo curso e o índice de evasão. Achados cruciais desse trabalho incluem a demonstração de que a taxa de evasão no primeiro ano de curso chega a ser de duas a três vezes maior do que nos anos subsequentes, e a desmistificação da

premissa de que fatores puramente financeiros seriam os únicos causadores do abandono; embora frequentemente declarados pelos discentes, os fatores de ordem acadêmica, as expectativas frustradas e a falta de integração inicial com a instituição constituem os reais motivadores do desestímulo.

Em relação ao uso de AM para abordar a evasão escolar, observa-se uma prevalência de algoritmos baseados em árvores de decisão, enquanto o nível universitário e a modalidade presencial constituem cenários frequentemente abordados [Jesus and Gusmão 2024]. A limitação de dados e a complexidade em medir os fatores que influenciam o desempenho dos alunos são desafios frequentemente encontrados [Magalhães et al. 2025]. Ao mesmo tempo, a ausência de fundamentação em teorias de aprendizagem e em conceitos de explicabilidade reduz a confiança nos resultados de modelos de predição de desempenho estudantil, restringindo sua adoção prática e intervenções preventivas direcionadas [Xiao et al. 2022].

Com o advento da pandemia de COVID-19 e a consequente transição para o ensino remoto emergencial, novas variáveis comportamentais foram introduzidas na literatura educacional, decorrentes da ampla adoção de Ambientes Virtuais de Aprendizagem. Neste contexto, trabalhos também analisaram os padrões-chave de evasão e como esses padrões se comparam antes e durante a pandemia [Colpo et al. 2024]. Algumas abordagens relacionam conceitos de interpretabilidade e justiça algorítmica no contexto da MDE, permitindo diferenciar modelos preditivos sob esses aspectos [Carvalho et al. 2025]. Por outro lado, também são avaliados os impactos estritamente técnicos das etapas de engenharia de dados sobre as explicações geradas, analisando como a aplicação de técnicas de balanceamento de dados pode impactar a geração de explicações de sucesso estudantil [Cavus and Kuzilek 2024].

Conforme revisado extensivamente por [Gama et al. 2014] e expandido por [Lu et al. 2019], um dos pressupostos mais desafiadores na implantação de modelos preditivos no mundo real é a natureza não-estacionária dos dados. Assim, o desvio de conceito (*concept drift*) refere-se a essas mudanças nas propriedades estatísticas das variáveis ao longo do tempo. De acordo com os autores, ignorar essa evolução temporal faz com que padrões aprendidos no passado percam a validade, degradando severamente a performance do modelo após sua implantação. Ambos os estudos citados detalham estratégias matemáticas para detecção e adaptação ao *drift*, recomendando a atualização contínua dos modelos preditivos por meio do reprocessamento de novas janelas temporais de dados.

Embora a literatura sobre *concept drift* esteja consolidada, revisões clássicas como as de [Gama et al. 2014] e [Lu et al. 2019] focam prioritariamente no monitoramento do desempenho preditivo. Recentemente, novos trabalhos buscam unir a detecção de desvios à interpretabilidade. É o caso de [Neto et al. 2025], que utiliza Redes de Função de Base Radial (RBFN) e Cadeias de Markov para mapear visual e probabilisticamente como os conceitos mudam ao longo do tempo.

Nessa mesma linha, outros estudos aplicam o método SHAP para explicar os desvios. O *framework* CDLEEDS, de [Haug et al. 2022], detecta mudanças locais para garantir que a importância das características (valores SHAP) permaneça válida e atualizada em fluxos de dados dinâmicos. Já [Decker et al. 2024] propõe o *framework* XPE, que utiliza Transporte Ótimo para alinhar distribuições e Valores de Shapley para quantificar

exatamente quais características são responsáveis pela queda de performance do modelo sob desvio.

O presente estudo propõe um avanço metodológico ao estender o monitoramento do desvio de conceito ao campo da interpretabilidade no contexto da predição de evasão de estudantes no ensino superior. Ao treinar algoritmos em conjuntos de dados progressivos, simulando uma evolução temporal, esta pesquisa avalia não apenas a evolução das métricas quantitativas, mas também investiga se o desvio estatístico recalibra a importância dos fatores explicativos calculados pelo SHAP, permitindo compreender de que forma o *concept drift* impacta decisões em ambientes educacionais.

3. Materiais e métodos

Para a realização dos experimentos, foi utilizada a plataforma *Google Colaboratory* e a linguagem de programação *Python*, comumente empregada para trabalhar com o AM, utilizando as bibliotecas *NumPy* e *Pandas* para o processamento dos dados, o que garante a preservação semântica ao longo do pipeline e o rigor matricial nas operações de álgebra linear. Para as classificações e métricas dos modelos, foram utilizados o *scikit-learn* e pacotes específicos, como o *XGBoost*, com o objetivo de estabelecer linhas de base transparentes para as avaliações.

3.1. Dataset

Para compor o conjunto de dados utilizado no experimento, foram utilizadas informações do Sistema Integrado de Gestão da Universidade Federal de Pelotas (UFPEL), abrangendo, em sua totalidade, 17.866 instâncias de alunos de graduação de todos os cursos, ingressantes entre 2014 e 2017, que, até o semestre de 2025/1, já estavam formados ou haviam desistido da graduação. Os atributos foram organizados em informações sociodemográficas e dados de desempenho acadêmico, coletados no primeiro semestre de cada discente, conforme descrito na Tabela 1. A variável alvo (Situação) foi definida de forma binária (“Formado”, “Evadido”), caracterizando, assim, um problema de classificação.

Durante o pré-processamento, foi utilizada a ferramenta *Ydata Profiling*, que permitiu a análise de cada variável, destacando fatores como a quantidade de dados faltantes, o desbalanceamento e as correlações entre eles, possibilitando, assim, a identificação de registros incoerentes. A partir dessa primeira análise, foram identificadas e removidas 519 linhas duplicadas ou com inconsistências nos atributos, incluindo valores negativos e observações consideradas implausíveis, resultando em um *dataset* com 17.347 instâncias.

3.2. Evolução temporal

Para simular a evolução temporal, o *dataset* foi dividido em quatro conjuntos de dados progressivos, cada um incorporando semestres adicionais em relação ao anterior, garantindo um acompanhamento dos parâmetros de decisão dos modelos em relação à evolução real dos fluxos de dados educacionais. Os *datasets* resultantes são descritos a seguir:

- *Dataset 1* - inclui alunos ingressantes de 2014-1 até 2016-1;
- *Dataset 2* - inclui alunos ingressantes de 2014-1 até 2016-2;
- *Dataset 3* - inclui alunos ingressantes de 2014-1 até 2017-1;
- *Dataset 4* - inclui alunos ingressantes de 2014-1 até 2017-2.

Tabela 1. Lista de atributos utilizados na análise após filtragem

Atributo
Idade de Ingresso
Sexo
Cor
Estado civil
Código do SISU
Escola Pública
Nível Superior
Turno
Tempo entre ensino médio e Faculdade
Área fundamental
Número de disciplinas
Número de créditos
Média do semestre
Número de créditos aprovados
Número de créditos reprovados
Número de créditos infrequentes
Número de créditos trancados
Número de exames

Ao realizar essa separação, poderá ser analisado de forma clara e objetiva como o crescimento na quantidade de dados afeta a capacidade de predição dos modelos, bem como o impacto na interpretabilidade.

3.3. Modelos

Para a obtenção dos modelos, cada *dataset* foi particionado em dois conjuntos: treino (70%) e teste (30%), de modo a manter uma amostra suficiente para seu treinamento e outra para posterior avaliação.

O tratamento de variáveis categóricas, como sexo e cor, foi realizado com a aplicação da técnica de *One-Hot Encoding* que, para cada categoria única em uma variável, cria uma nova coluna binária independente (“1” para indicar a presença do atributo e “0” para indicar a ausência).

Além disso, para tratar o desbalanceamento das classes, empregou-se a técnica SMOTE (*Synthetic Minority Over-sampling Technique*), aplicada exclusivamente ao conjunto de treino, que gera dados sintéticos a partir dos “k” vizinhos mais próximos das classes minoritárias. Assim, essa técnica é mais eficiente do que o *oversampling* tradicional, que simplesmente duplica instâncias da classe minoritária. Com a aplicação dessa estratégia, o conjunto de treinamento foi perfeitamente balanceado, resultando em uma distribuição simétrica de 50% das instâncias para a classe majoritária e 50% para a classe minoritária, mitigando o viés do algoritmo em direção ao comportamento predominante.

Para o AM, foram empregadas quatro técnicas:

- *Random Forest* - Constrói múltiplas árvores de decisão independentes, permitindo o cálculo da importância das variáveis por meio da diminuição média da impureza em todos os nós que utilizaram certa variável;
- *Gradient Boosting* - Constrói as árvores de forma sequencial, com cada árvore subsequente tentando corrigir os erros residuais deixados pela árvore anterior;

- *XGBoost* - Evolução do algoritmo de Gradient Boosting, processando de forma paralela e possuindo termos de regularização (L1 e L2) que penalizam modelos excessivamente complexos;
- *ANN-MLP* - Rede neural composta por uma camada de entrada, camadas intermediárias ocultas e uma camada de saída, conectadas por pesos sinápticos.

A escolha de múltiplos modelos foi feita com o objetivo de comparar diferentes abordagens de modelagem. Para cada um dos quatro *datasets* especificados anteriormente, foram treinados os algoritmos utilizando os 4 métodos, totalizando 16 modelos distintos. Por fim, cada um desses modelos foi avaliado no conjunto de teste mais atualizado.

4. Resultados

Para avaliar o desempenho preditivo dos modelos, foram utilizadas as seguintes métricas consagradas:

- Área sob a curva ROC (AUC-ROC) - Avalia a performance em todos os limiares de decisão, plotando a sensibilidade em função da proporção de falsos alarmes;
- *Accuracy* - Representa a proporção de predições corretas em relação ao total de predições realizadas;
- *Precision* - Quantifica a proporção de identificações positivas que foram realmente positivas, penalizando a ocorrência de falsos positivos.

A análise dos resultados quantitativos evidencia uma tendência consistente de aprimoramento do desempenho à medida que os modelos são atualizados com novos dados. Conforme ilustrado na tabela de resultados, o benefício de um conjunto de dados mais amplo e diverso mostra-se inequívoco para um melhor desempenho. Para fins de concisão e para destacar os extremos do desvio de conceito, a Tabela 2 apresenta apenas os resultados referentes aos conjuntos inicial e final

Tabela 2. Tabela de resultados com os modelos treinados.

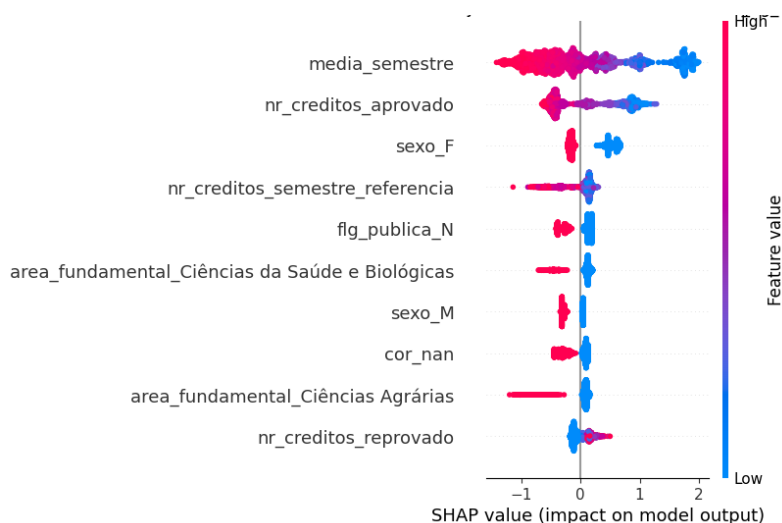
<i>Dataset</i>	<i>Modelo</i>	<i>Evaluated on Test Set</i>	<i>Accuracy</i>	<i>Precision</i>	<i>AUC-ROC</i>
Treino 1	RandomForest	Teste 4	0.777157	0.779842	0.857363
Treino 1	XGBoost	Teste 4	0.767572	0.773823	0.854305
Treino 1	ANN_MLP	Teste 4	0.786741	0.800438	0.856805
Treino 1	GradientBoosting	Teste 4	0.767572	0.773823	0.854305
Treino 4	RandomForest	Teste 4	0.783546	0.784828	0.866156
Treino 4	XGBoost	Teste 4	0.783546	0.790780	0.871143
Treino 4	ANN_MLP	Teste 4	0.777955	0.782862	0.860193
Treino 4	GradientBoosting	Teste 4	0.783546	0.790780	0.871143

Por exemplo, ao analisar o modelo *XGBoost*, pode se observar uma melhora na métrica AUC-ROC de 0,854 (quando treinado com o *Dataset 1*) para 0,871 (quando treinado com o *Dataset 4*), evidenciando essa melhora devido à progressão temporal dos dados. Esse padrão de melhora/evolução foi reproduzido nos demais modelos, indicando que a maior heterogeneidade dos perfis estudantis contribui para a construção de modelos mais generalistas e precisos. Essas evidências sugerem que sistemas de predição de evasão se beneficiam de treinamentos contínuos com dados atualizados, fornecendo

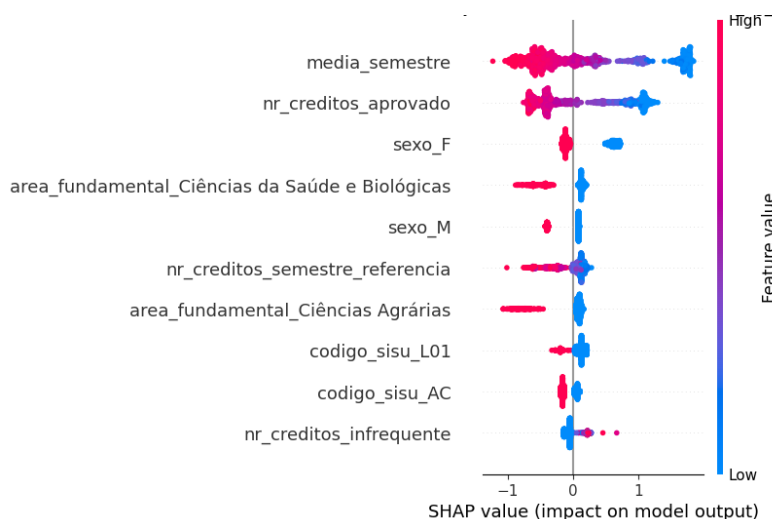
subsídios iniciais para que possamos responder a questão de pesquisa Q1 “A atualização contínua dos dados melhora o desempenho preditivo?”.

Para a análise da interpretabilidade, foi aplicada a ferramenta *SHAP* (SHapley Additive exPlanations), uma abordagem que calcula a contribuição marginal exata de cada variável para a predição final do modelo que apresentou os melhores resultados (*XGBoost*), treinado tanto com o *dataset 1* quanto com o *dataset 4*.

Figura 1. SHAP do modelo XGBoost com os 10 atributos mais relevantes.



(a) Modelo obtido com o *dataset 1*.



(b) Modelo obtido com o *dataset 4*.

Neste gráfico, o eixo vertical ordena as variáveis de cima para baixo conforme sua importância global para as decisões do modelo. O eixo horizontal representa o valor SHAP, que quantifica o impacto de cada característica na predição final, valores positivos (à direita do centro) indicam um aumento na propensão à classe alvo, enquanto valores negativos (à esquerda) indicam uma redução. Por fim, a escala de cores ilustra o valor original da variável para cada observação, variando de azul (valores baixos) a vermelho (valores altos), permitindo identificar como a magnitude de uma variável direciona a

predição. Essa análise evidenciou que as variáveis de maior relevância estão fortemente associadas ao desempenho acadêmico inicial do aluno. Atributos como “média do semestre” e “número de créditos aprovados” destacaram-se como os principais preditores de evasão em ambos os modelos. Esses resultados trazem avanços consideráveis para que possamos responder os questionamentos propostos pelas questões de pesquisa Q2 “De que forma a atualização dos dados altera os fatores explicativos identificados pelo SHAP?” e Q3 “É possível detectar indícios de *concept drift* por meio das explicações?”.

Esse achado é relevante, pois reforça a hipótese de que o engajamento e o êxito do estudante no primeiro semestre constituem fatores críticos para sua permanência. Assim, uma média baixa ou um número reduzido de aprovações pode servir como um alerta precoce para as equipes pedagógicas, possibilitando intervenções direcionadas e fundamentadas em evidências.

Ao comparar esses dois resultados, percebe-se que a evolução do *dataset* não alterou os atributos mais dominantes (Média do semestre e número de créditos aprovados). Porém, nota-se uma alteração na influência de outros atributos, evidenciando a ocorrência do fenômeno de *concept drift* (desvio de conceito). Essa mudança indica que a relação subjacente entre as características secundárias dos discentes e a variável alvo se altera ao longo do tempo. Por exemplo, os atributos “número de créditos no semestre de referência” e “escola pública” possuem peso maior no primeiro modelo (Figura 1a), enquanto os atributos “Área fundamental Ciências da saúde e biológicas” e “Área fundamental Ciências Agrárias” mostraram-se mais importantes no segundo modelo (Figura 1b), refletindo uma recalibração e uma tendência ao aperfeiçoamento de sua predição, visto a melhoria em suas métricas.

5. Conclusão

Este estudo evidenciou a relevância do uso de técnicas de interpretabilidade para um melhor entendimento de algoritmos de AM, principalmente no campo acadêmico e na evasão dos discentes, além de destacar o uso de conjuntos de dados temporalmente evolutivos como uma ferramenta de aprimoramento dessas técnicas. Constatou-se que a incorporação de dados de semestres mais recentes contribui para a consistência das predições realizadas pelos modelos, tornando-os mais robustos e precisos. Ademais, por meio da aplicação do método *SHAP*, foi possível identificar com mais clareza os fatores críticos do desempenho acadêmico que permitem prever o risco de evasão desde o primeiro semestre na universidade. Com estes resultados, será consolidada a discussão sobre o uso transparente da inteligência artificial na educação, uma vez que a união de dados evolutivos e técnicas de explicabilidade contribuirá para que sinais de alerta cheguem às coordenações de forma fundamentada e compreensível, promovendo intervenções pedagógicas direcionadas e um uso mais responsável e ético da Inteligência Artificial no ensino.

Como desdobramentos futuros e próximos passos deste estudo, planeja-se ampliar a análise dos resultados parciais aqui apresentados, como a melhoria preditiva e recalibração das variáveis frente ao *concept drift*, viabilizando a aplicação prática em instituições de ensino. Adicionalmente, a pesquisa avançará na validação metodológica, comparando as explicações extraídas com outras ferramentas de interpretabilidade, como o *LIME*, visando a consolidação de um framework robusto para a gestão educacional

Declaração sobre uso de inteligência artificial

Em conformidade com o Código de Conduta para autores da SBC, declara-se o uso de ferramentas de inteligência Artificial Generativa apenas para auxiliar na escrita e na revisão do texto. O planejamento da pesquisa, a execução dos experimentos, a análise de dados e a elaboração das conclusões foram realizados integralmente pelos pesquisadores. Os autores assumem a responsabilidade sobre todo o conteúdo do artigo submetido.

Referências

- Baker, R. et al. (2010). Data mining for education. *International encyclopedia of education*, 7(3):112–118.
- Bakhshinategh, B., Zaiane, O. R., ElAtia, S., and Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, 23:537–553.
- Carvalho, C. S., de Mattos, J. C. B., and de Aguiar, M. S. (2025). Lime-ec: an agnostic approach to quantifying model interpretability using explanation clustering. In *2025 12th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, pages 549–555.
- Carvalho, C. S., Mattos, J. C. B. d., and Aguiar, M. S. d. (2026). Model interpretability in the educational data mining context: A systematic literature mapping. *Brazilian Journal of Computers in Education*, 34:314–355.
- Carvalho, D. V., Pereira, E. M., and Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, 8(8).
- Cavus, M. and Kuzilek, J. (2024). An effect analysis of the balancing techniques on the counterfactual explanations of student success prediction models. *Journal of Measurement and Evaluation in Education and Psychology*, 15:302 – 317.
- Colpo, M. P., Primo, T. T., and de Aguiar, M. S. (2024). Lessons learned from the student dropout patterns on covid-19 pandemic: An analysis supported by machine learning. *British Journal of Educational Technology*, 55(2):560 – 585.
- Decker, T., Koebler, A., Lebacher, M., Thon, I., Tresp, V., and Buettner, F. (2024). Explanatory model monitoring to understand the effects of feature shifts on performance. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, page 550–561, New York, NY, USA. Association for Computing Machinery.
- Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4):1–37.
- Haug, J., Braun, A., Zürn, S., and Kasneci, G. (2022). Change detection for local explainability in evolving data streams. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, page 706–716, New York, NY, USA. Association for Computing Machinery.
- Jesus, J. A. d. and Gusmão, R. P. d. (2024). Investigação da evasão estudantil por meio da mineração de dados e aprendizagem de máquina: Um mapeamento sistemático. *Revista Brasileira de Informática na Educação*, 32:807–841.

- Linardatos, P., Papastefanopoulos, V., and Kotsiantis, S. (2021). Explainable ai: A review of machine learning interpretability methods. *Entropy*, 23(1).
- Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., and Zhang, G. (2019). Learning under concept drift: A review. *IEEE Transactions on Knowledge and Data Engineering*, 31(12):2346–2363.
- Magalhães, P. V. S., Sousa, R. A. d. S., Santos, A. C. B., Aguiar, F. J. M., and Milfont, R. T. P. (2025). Modelos de inteligência artificial para predição de evasão escolar: Uma revisão sistemática. *Revista Eletrônica de Iniciação Científica em Computação*, 23(1):48–54.
- Molnar, C. (2022). *Interpretable Machine Learning*. 2 edition.
- Neto, R., Alencar, B., Gomes, H. M., Bifet, A., Gama, J., Cassales, G., and Rios, R. (2025). Rmddm: an unsupervised and interpretable concept drift detection method for data streams. *Data Mining and Knowledge Discovery*, 39(6):85.
- Saxena, N. A., Huang, K., DeFilippis, E., Radanovic, G., Parkes, D. C., and Liu, Y. (2019). How do fairness definitions fare? examining public attitudes towards algorithmic definitions of fairness. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 99–106.
- Silva Filho, R. L. L. e., Motejunas, P. R., Hipólito, O., and Lobo, M. B. d. C. M. (2007). A evasão no ensino superior brasileiro. *Cadernos de Pesquisa*, 37(132):641–659.
- Xiao, W., Ji, P., and Hu, J. (2022). A survey on educational data mining methods used for predicting students' performance. *Engineering Reports*, 4(5):e12482.