

Modelagem Preditiva Explicável para Apoio à Prevenção da Evasão em Cursos de Graduação

Ronald César C. da Silva¹, Francisco Glaubos N. Climaco¹,
Raphaella Abreu C. C. Moreira¹, Jaiver E. J. Figueroa¹, Tiago B. Borchart¹

¹ Universidade Federal do Maranhão (UFMA)

São Luís – MA – Brasil

{ronald.cesar, francisco.glaubos, raphaella.moreira,
jaiver.figueroa, tiago.bonini}@ufma.br

Abstract. *This study investigates academic dropout at a Brazilian public university through the integration of predictive modeling and Explainable Artificial Intelligence (XAI) techniques. The dataset comprises 5,876 undergraduate students across four degree programs, considering institutional performance indicators and demographic variables. Classification models were evaluated using AUC-PR, and global and local explainability methods were applied to the selected model in each program. Results indicate that academic efficiency indicators (IECH and IEPL) and the completion average (MC) are the most influential predictors of persistence or dropout, whereas demographic factors play a secondary role. Course-level analysis reveals distinct risk patterns, highlighting the need for context-specific intervention strategies. The proposed framework enhances model transparency and provides methodological support for institutional dropout prevention planning.*

Keywords: *Artificial Intelligence; Academic Dropout; Predictive Modeling; Explainable AI; SHAP.*

Resumo. *Este estudo investiga a evasão acadêmica em uma universidade pública brasileira por meio da integração entre modelagem preditiva e técnicas de Inteligência Artificial Explicável (XAI). A amostra compreende 5.876 estudantes de quatro cursos de graduação, considerando indicadores institucionais de desempenho e variáveis demográficas. Modelos de classificação foram avaliados com AUC-PR e, sobre o modelo selecionado em cada curso, aplicaram-se métodos de explicabilidade global e local. Os resultados indicam que os índices de eficiência acadêmica (IECH e IEPL) e a média de conclusão (MC) são os principais preditores da permanência ou evasão, enquanto fatores demográficos apresentam menor influência. A análise por curso evidencia padrões distintos de risco, reforçando a necessidade de estratégias específicas por área. O framework proposto amplia a transparência dos modelos e oferece subsídios metodológicos para o planejamento de ações institucionais de prevenção à evasão.*

Palavras-chave: *Inteligência Artificial; Evasão Acadêmica; Modelagem Preditiva; XAI; SHAP.*

1. Introdução

A evasão acadêmica permanece um dos desafios mais persistentes do ensino superior, com impactos que se estendem para além do campo educacional, alcançando dimensões sociais

e econômicas mais amplas. Relatórios da Organisation for Economic Co-operation and Development [OECD 2023] indicam que a conclusão de cursos de graduação permanece desigual entre países e grupos sociais. No Brasil, o Instituto Semesp [Semesp 2025] estima que a taxa de evasão em cursos presenciais de universidades públicas atingiu 23,1% em 2023, gerando perdas financeiras significativas e fragilizando políticas de democratização do acesso e da permanência na graduação. Padrões semelhantes têm sido observados em diferentes contextos, o que reforça o caráter sistêmico do fenômeno e a necessidade de estratégias analíticas que permitam o monitoramento contínuo.

Além das consequências institucionais, a não conclusão do ensino superior está associada a efeitos pessoais de longo prazo, como menor empregabilidade e redução do bem-estar pessoal, conforme discutem [Neugebauer et al. 2025]. Nesse cenário, torna-se crucial dispor de instrumentos que identifiquem precocemente estudantes em risco, de modo a orientar intervenções pedagógicas e ações de apoio mais bem direcionadas.

Nos últimos anos, a Inteligência Artificial (IA), por meio de técnicas de *Machine Learning* (ML), tem demonstrado potencial para apoiar a gestão acadêmica. Estudos como os de [Pérez et al. 2018], [Ulloa-Cázar et al. 2021], [Niyogisubizo et al. 2022] e [Kocsis and Molnár 2025] empregam árvores de decisão, redes neurais e métodos de *ensemble* para estimar o risco de desligamento com base em informações acadêmicas e demográficas. Apesar do bom desempenho preditivo, diferentes revisões apontam que a ausência de mecanismos explicativos transparentes limita a adoção prática desses modelos em instituições de ensino, nas quais interpretabilidade, justiça e confiabilidade são requisitos centrais para o uso responsável de IA [Tete et al. 2022], [Rodrigues et al. 2024] e [Nagy and Molontay 2024].

Nesse contexto, métodos de Inteligência Artificial Explicável (XAI), como *SHAP* e *LIME*, têm sido explorados para aproximar modelos preditivos do cotidiano institucional, permitindo compreender fatores que influenciam o risco de evasão e apoiar decisões pedagógicas [Nagy and Molontay 2024].

Este estudo aplica técnicas de XAI a modelos preditivos de evasão construídos com dados de cursos presenciais de uma universidade pública brasileira. O objetivo é identificar os fatores que mais influenciam as decisões dos modelos e avaliar como as explicações produzidas podem apoiar o planejamento de estratégias de prevenção à evasão. Para isso, propõe-se um processo analítico composto por modelagem segmentada por curso, integração de métodos globais e locais de explicabilidade e interpretação contextualizada dos resultados, estruturado de modo a ser reprodutível e adaptável a outras instituições.

A pesquisa também dialoga com lacunas da literatura, que ainda carece de estudos que explicitem como variáveis influenciam as previsões e como tais explicações podem ser traduzidas em ações pedagógicas concretas. Com base nessas lacunas, este estudo é guiado pelas seguintes questões de pesquisa: **RQ1** – Quais variáveis exercem maior influência sobre as decisões dos modelos preditivos de evasão? **RQ2** – De que maneira as técnicas de XAI ajudam a identificar padrões de desempenho associados ao risco de evasão? **RQ3** – Como as explicações podem apoiar os gestores no planejamento de intervenções pedagógicas voltadas à prevenção da evasão?

Este artigo está organizado da seguinte forma. A **Seção 2** revisa trabalhos relacionados e o referencial teórico sobre modelagem preditiva e explicabilidade na educação.

A **Seção 3** descreve o conjunto de dados e os procedimentos metodológicos adotados. A **Seção 4** apresenta os resultados da modelagem e das análises de explicabilidade. **Seção 5** conclui o artigo, sintetizando as principais contribuições, limitações e direções para pesquisas futuras.

2. Trabalhos relacionados

A Inteligência Artificial Explicável (XAI) busca ampliar a transparência dos modelos de aprendizado de máquina, permitindo que suas decisões sejam analisadas de forma compreensível por diferentes públicos [Klockmann et al. 2022]. No contexto educacional, a interpretabilidade é central para o uso ético de previsões, especialmente na identificação de estudantes em risco de evasão e na mitigação de vieses, o que favorece a adoção responsável de tecnologias de IA na gestão acadêmica [Melo et al. 2022] e [da Silva et al. 2022].

Entre os métodos *post hoc* mais utilizados em XAI destacam-se o *Local Interpretable Model-Agnostic Explanations* (LIME), o *SHapley Additive exPlanations* (SHAP), o *Anchors* e o *Permutation Importance* (PI). Em conjunto, essas técnicas oferecem perspectivas complementares: SHAP e LIME produzem explicações quantitativas em níveis local e global, Anchors resume decisões em regras de fácil interpretação e o PI sintetiza a importância relativa das variáveis [Ribeiro et al. 2016], [Lundberg and Lee 2017], [Ribeiro et al. 2018], [Molnar 2025] e [Saarela and Podgorelec 2024]. Essa combinação tem se mostrado promissora para tornar modelos complexos mais transparentes em aplicações educacionais.

A previsão de evasão por meio de *Machine Learning* (ML) remonta às primeiras iniciativas de mineração de dados educacionais. Estudos como [Heredia et al. 2015] e [Hernández-Gonzalez et al. 2016] demonstraram que modelos interpretáveis, como árvores de decisão e regressão logística, podem alcançar bom desempenho ao estimar a probabilidade de desligamento, mantendo clareza na identificação de fatores associados à evasão. Pesquisas mais recentes incorporam princípios de XAI em contextos educacionais, indicando que explicações adequadas ao perfil de gestores, docentes e estudantes tendem a aumentar a confiança e a aceitação dos modelos [Fiok et al. 2022] e [Farrow 2023].

No contexto da evasão no ensino superior, diferentes estudos têm explorado o uso de XAI como apoio à permanência estudantil. [da Silva et al. 2024] compararam *SHAP*, *LIME* e *Anchors*, indicando maior estabilidade do *SHAP* e maior acessibilidade interpretativa do *Anchors*. [Barbosa et al. 2024] propuseram um painel interativo baseado em SHAP para visualização de fatores de risco e suas contribuições individuais, enquanto [Saqr and López-Pernas 2024] destacam o potencial de abordagens personalizadas para orientar recomendações a estudantes com alto risco de desligamento.

Apesar desses avanços, permanecem lacunas importantes. Muitos estudos se concentram apenas em identificar estudantes em risco, sem detalhar como as explicações podem ser integradas a estratégias concretas de intervenção pedagógica. Além disso, é comum tratar a evasão de forma agregada, desconsiderando as especificidades de cada curso ou área de formação. Sob essa perspectiva, o presente estudo se distingue ao integrar múltiplas técnicas de XAI em um processo analítico reutilizável que combina modelagem preditiva segmentada por curso, explicações globais e locais e interpretação contextual dos resultados. O objetivo é oferecer um framework de interpretabilidade replicável, capaz de gerar evidências acionáveis para apoiar decisões institucionais e orientar estratégias de

identificação e mitigação do risco de evasão.

3. Metodologia

O estudo adota uma abordagem aplicada, organizada em um processo analítico que abrange desde a extração dos dados até a explicação das predições, com foco em apoiar decisões educacionais voltadas à prevenção da evasão. O procedimento foi estruturado de forma modular para facilitar sua replicação em outros contextos institucionais.

O processo metodológico é composto por quatro etapas: (i) *Compreensão dos dados*, com verificações de integridade e exclusão de registros incompletos; (ii) *Preparação dos dados*, envolvendo codificação, normalização e balanceamento das classes; (iii) *Modelagem*, conduzida separadamente para cada curso; e (iv) *Explicabilidade*, na qual *SHAP*, *PI*, *LIME* e *Anchors* são aplicados para gerar interpretações globais e locais. A Figura 1 sintetiza esse fluxo, articulando o pré-processamento, a indução dos modelos e a análise explicável ao uso prático dos resultados.

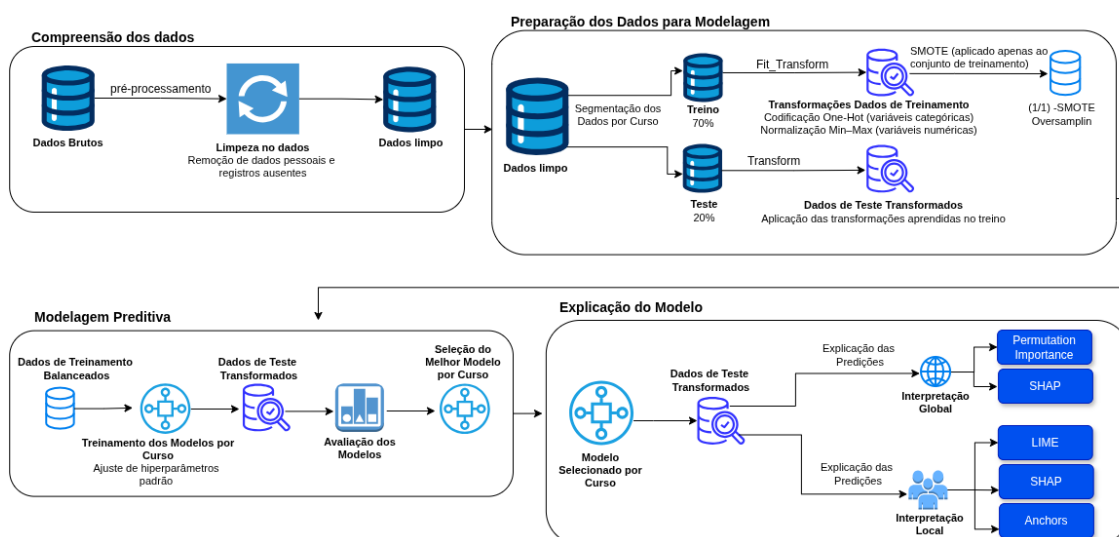


Figura 1. Fluxo metodológico do processo de interpretabilidade.

3.1. Compreensão dos Dados

A base de dados reúne 5.876 estudantes de quatro cursos presenciais — Enfermagem (ENF), Odontologia (ODO), Ciência da Computação (CCO) e Bacharelado Interdisciplinar em Ciência e Tecnologia (BICT) — com ingresso entre 2011 e 2024 em uma universidade pública brasileira. Os cursos foram selecionados por combinarem perfis acadêmicos distintos e registros institucionais consistentes, o que permite avaliar a abordagem proposta em contextos diferentes.

Os dados foram extraídos do Sistema Integrado de Gestão de Atividades Acadêmicas e submetidos a procedimentos de consistência. Registros com valores ausentes em variáveis essenciais (*Status*, *IECH*, *IEPL* e *MC*) foram removidos, o que totalizou cerca de 4% da amostra. Informações sensíveis foram excluídas conforme a LGPD (Lei nº 13.709/2018), assegurando tratamento ético e confidencial dos dados.

As variáveis utilizadas, detalhadas na Tabela 1, incluem indicadores de desempenho acadêmico (*MC*, *IECH* e *IEPL*) e atributos demográficos. Os indicadores de desempenho

são calculados pelo sistema acadêmico a partir de fórmulas institucionais padronizadas e representam dimensões complementares da trajetória dos estudantes:

- **Média de Conclusão (MC):** média ponderada das notas finais nas disciplinas aprovadas, sintetizando o desempenho global.
- **Índice de Eficiência em Carga Horária (IECH):** proporção da carga horária matriculada que resultou em aprovação, refletindo a eficiência na conversão de créditos cursados em créditos concluídos.
- **Índice de Eficiência em Períodos Letivos (IEPL):** razão entre a carga horária acumulada e a carga horária esperada para o número de semestres cursados, indicando o ritmo de progressão em relação ao tempo ideal do curso.

Tabela 1. Descrição das Variáveis utilizadas neste estudo.

Variável	Descrição
Curso	Curso de graduação no qual o estudante está matriculado.
Sexo	Gênero registrado no sistema institucional de informações.
Tip_Red_Ensino	Tipo de escola frequentada durante o ensino médio.
Tipo_Raça	Identificação racial ou étnica autodeclarada.
Média_de_Conclusão(MC)	Média ponderada das notas obtidas nas disciplinas concluídas.
Índice_de_Eficiência_em_Carga_Horária(IECH)	Eficiência da carga horária matriculada que foi efetivamente convertida em aprovações.
Índice_de_Eficiência_em_Períodos_Letivos(IEPL)	Ritmo de progressão acadêmica em relação ao tempo ideal do curso.
Idade_Ingresso	Idade do estudante no momento do ingresso na universidade.
Região_Origem	Local de origem do estudante (capital, outros municípios do estado ou de outros estados).
Forma_Ingresso	Modalidade de ingresso (ex.: categoria de seleção do SISU).
Status	Situação acadêmica final: Concluído ou Desligado (evasão).

A Figura 2 apresenta a distribuição da situação final dos estudantes, considerando todos os ingressantes entre 2011 e 2024, por curso, e evidencia contrastes marcantes: CCO e BICT concentram as maiores proporções de desligamento, enquanto ENF e ODO exibem distribuições mais equilibradas entre concluintes e desligados.

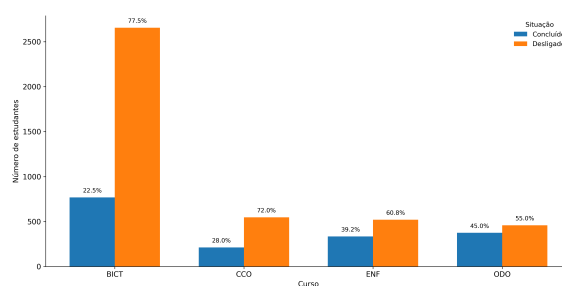


Figura 2. Situação final dos estudantes por curso (2011–2024)

A análise exploratória das correlações entre os indicadores quantitativos, ilustrada na Figura 3, revelou uma associação forte entre *IECH* e *IEPL*, o que confirma a complementaridade entre eficiência e ritmo de progressão. O indicador *MC* também apresentou correlação positiva com ambos, sugerindo convergência entre o desempenho acadêmico e o avanço curricular. Em relação ao desfecho (*Status*), observou-se correlação moderada de *IECH* e *IEPL* com a conclusão do curso, mais elevada em BICT e CCO e menor em ODO e ENF, enquanto a idade de ingresso apresentou baixa associação com o risco de desligamento.

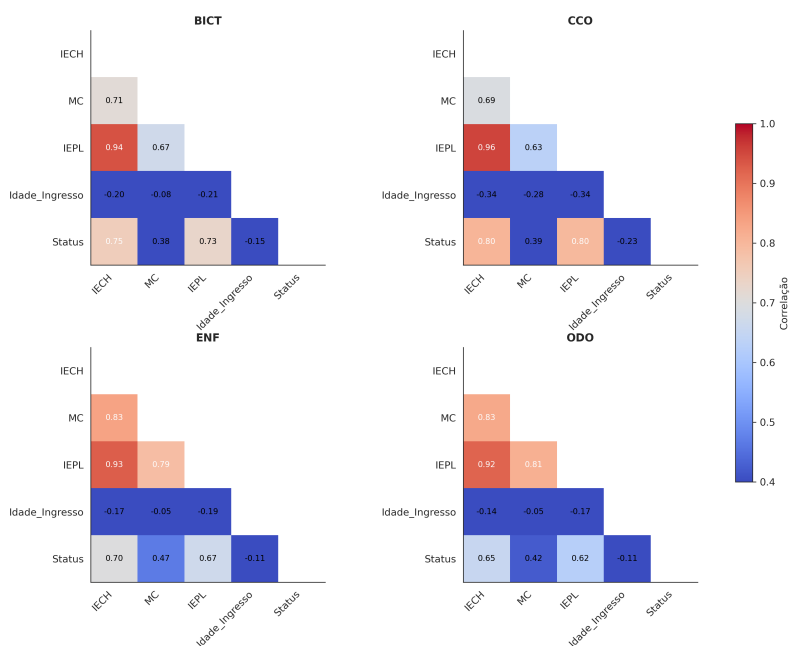


Figura 3. Matrizes de correlação entre indicadores acadêmicos, idade de ingresso e situação final dos estudantes, segmentadas por curso.

Esses resultados evidenciam redundâncias parciais entre os preditores e diferenças estruturais entre os cursos, reforçando a opção metodológica por uma modelagem explicável segmentada por área de formação.

3.2. Preparação dos Dados

Todas as transformações foram ajustadas exclusivamente no conjunto de treinamento e posteriormente aplicadas ao conjunto de teste por meio de *pipelines* do Scikit-learn, reduzindo o risco de vazamento de informação e favorecendo a reprodutibilidade [Géron 2021]. A amostra foi particionada em 70% para treino e 30% para teste, com estratificação por curso e preservação da proporção entre concluintes e desligados, assegurando consistência na avaliação fora da amostra.

As variáveis categóricas foram codificadas por *One-Hot Encoding (OHE)* apenas no conjunto de treinamento, fixando-se as categorias observadas nessa etapa e replicando-as no conjunto de teste. O uso de OHE evita ordens artificiais em atributos nominais e preserva a interpretabilidade ao representar cada categoria por um indicador binário [Harrison 2020]. As variáveis contínuas, por sua vez, foram normalizadas com *Min-Max Scaling*, a fim de harmonizar as escalas e reduzir o efeito de magnitudes muito distintas sobre o ajuste dos modelos.

Devido ao desbalanceamento da variável-alvo (*Status*) em todos os cursos, foi utilizado o *SMOTE (Synthetic Minority Over-sampling Technique)* após as transformações e somente no conjunto de treinamento, com o objetivo de mitigar vieses na indução dos classificadores sem alterar a distribuição original dos dados de teste [da Silva Dantas et al. 2024]. Essa decisão contribui para uma avaliação mais justa do desempenho preditivo e para a construção de modelos mais estáveis.

3.3. Modelagem

A etapa de modelagem teve como objetivo construir modelos capazes de prever o risco de evasão e comparar seu desempenho em diferentes áreas de formação. Os modelos foram ajustados separadamente para cada curso, de modo a captar as particularidades nas trajetórias acadêmicas e nos padrões de desligamento observados em ENF, ODO, CCO e BICT.

Foram avaliados três algoritmos amplamente utilizados em tarefas preditivas complexas — *XGBoost*, *Random Forest* e *Multilayer Perceptron (MLP)* —, que costumam apresentar bom desempenho, embora com interpretabilidade direta limitada [Saarela and Podgorelec 2024]. Como linha de base interpretável, foi incluída a *Logistic Regression*, adotada como referência transparente por oferecer uma estrutura explicativa baseada em coeficientes [Marques et al. 2023].

Com o intuito de privilegiar a comparabilidade entre cursos e algoritmos, optou-se por não realizar ajuste fino de hiperparâmetros nem otimização de limiares de decisão. Todos os modelos foram executados com configurações padrão das bibliotecas utilizadas e limiar de classificação fixo em 0,5. Essa escolha evita que ajustes específicos de cada algoritmo ou base de dados influenciem a comparação e facilita a replicação do procedimento em outros contextos institucionais.

O desempenho foi avaliado pela Área Sob a Curva Precisão–Revocação (AUC–PR), métrica apropriada para classes desbalanceadas por refletir a relação entre precisão e revocação quando a classe positiva é minoritária [Saito and Rehmsmeier 2015]. A incerteza foi estimada por intervalos de confiança de 95% obtidos via reamostragem bootstrap com 1.000 iterações sobre as predições do conjunto de teste. Em cada curso, o modelo final foi selecionado pela maior média de AUC–PR, considerando os intervalos de confiança para evitar escolhas baseadas em diferenças instáveis.

3.4. Explicabilidade

A etapa de explicabilidade integrou a avaliação e a interpretação dos modelos, buscando que as predições fossem não apenas precisas, mas também compreensíveis para apoiar decisões educacionais relacionadas à prevenção da evasão. Neste estudo, adotaram-se duas perspectivas complementares — global e local — para compreender tanto o comportamento geral dos modelos quanto a lógica de predições individuais.

Na perspectiva global, o foco foi a importância relativa das variáveis. Utilizaram-se as técnicas *PI* e *SHAP*. A *PI* estimou a relevância de cada atributo a partir da variação na AUC–PR quando seus valores eram permutados aleatoriamente [Molnar 2025]. O *SHAP* quantificou a contribuição média das variáveis para as predições, permitindo identificar fatores mais influentes e possíveis interações [Lundberg and Lee 2017]. Em conjunto, essas abordagens forneceram uma visão interpretável dos principais determinantes do risco de evasão em cada curso.

Na perspectiva local, três métodos — *SHAP*, *LIME* e *Anchors* — foram aplicados para interpretar predições de estudantes específicos e avaliar a consistência das decisões. O *SHAP* decompôs a probabilidade predita em contribuições positivas ou negativas de cada variável, evidenciando quais atributos aumentaram ou reduziram o risco estimado. O *LIME* aproximou o comportamento do modelo em torno de exemplos selecionados

por meio de regressões lineares locais, destacando as variáveis mais influentes em cada caso [Ribeiro et al. 2016]. Já o Anchors derivou regras de decisão do tipo “se... então...”, traduzindo o comportamento do modelo em enunciados lógicos de fácil interpretação [Ribeiro et al. 2018]. A integração desses métodos permitiu comparar explicações entre perfis com diferentes níveis de risco e verificar a coerência das justificativas fornecidas pelos modelos.

4. Resultados e Discussão

4.1. Modelagem preditiva

O desempenho dos modelos foi avaliado por meio da AUC-PR, com intervalos de confiança de 95% obtidos a partir de 1.000 reamostragens *bootstrap*. De modo geral, os modelos baseados em árvores (*XGBoost* e *Random Forest*) apresentaram escores de AUC-PR mais elevados e menor dispersão, como sintetizado na Tabela 2.

Tabela 2. Desempenho dos Modelos Preditivos por Curso (AUC-PR com IC 95%).

Curso	Modelo	AUC-PR (95% CI)	Std. Dev.
ODO	XGBoost	96.15 [93.26 – 98.52]	0.013
	Random Forest	95.38 [91.64 – 98.50]	0.016
	MLP	86.35 [78.86 – 92.81]	0.034
	Logistic Regression	85.16 [78.62 – 91.94]	0.038
ENF	Random Forest	97.70 [95.84 – 99.03]	0.082
	XGBoost	97.58 [95.64 – 98.99]	0.081
	MLP	89.47 [82.55 – 94.90]	0.031
	Logistic Regression	88.31 [85.00 – 94.18]	0.034
BICT	Logistic Regression	92.23 [89.29 – 94.91]	0.014
	MLP	91.28 [87.95 – 94.27]	0.016
	XGBoost	89.83 [84.45 – 92.18]	0.024
	Random Forest	89.03 [82.13 – 92.97]	0.020
CCO	Random Forest	96.93 [93.93 – 99.10]	0.013
	XGBoost	95.97 [91.84 – 99.05]	0.018
	MLP	94.62 [88.68 – 98.60]	0.024
	Logistic Regression	93.88 [86.90 – 98.97]	0.031

No curso de Odontologia (ODO), o *XGBoost* obteve o melhor desempenho médio AUC-PR em torno de 0,96 e superou o *Random Forest*. Em Enfermagem (ENF), o padrão se inverteu: o *Random Forest* apresentou AUC-PR ligeiramente superior. No BICT, a *Logistic Regression* apresentou o melhor desempenho, com AUC-PR de 0,92, superando tanto o *MLP* quanto os modelos baseados em árvores, o que sugere que relações aproximadamente lineares foram suficientes para distinguir concluintes e desligados nesse contexto. Em Ciência da Computação (CCO), o *Random Forest* atingiu a maior AUC-PR (cerca de 0,97), consolidando-se como o modelo mais adequado para esse curso.

Esses resultados dialogam com revisões que apontam *XGBoost* e *Random Forest* como abordagens recorrentes e eficazes na previsão de evasão [Rodrigues et al. 2024], [da Silva et al. 2024], mas indicam que a escolha do algoritmo depende da estrutura curricular e dos padrões de desempenho de cada curso. Em termos práticos, isso reforça a relevância de estratégias de modelagem segmentada para apoiar a gestão acadêmica, pois perfis distintos de evasão demandam modelos alinhados às características de cada curso, produzindo indicadores de risco mais confiáveis para o planejamento de ações de permanência.

4.2. Explicabilidade global

A Figura 4 mostra um padrão recorrente entre os programas: *IECH* aparece como o preditor mais influente, seguido por *IEPL* e *MC*. Em contraste, atributos demográficos e de ingresso

— como gênero, raça, tipo de escola, modalidade de acesso e idade — desempenham papel secundário nas predições. Esse padrão reforça que o risco de evasão está mais associado ao desempenho e ao ritmo de progressão do que às características pessoais dos estudantes.

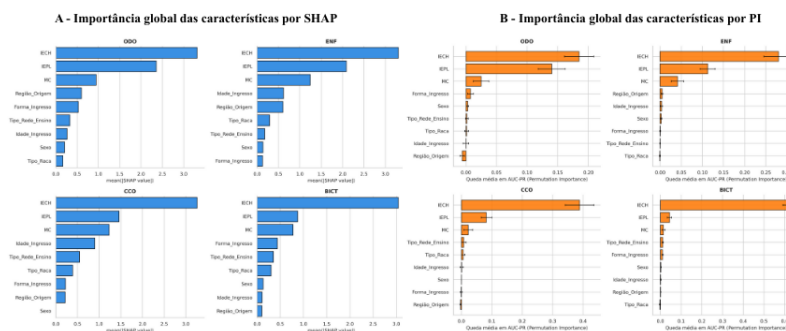


Figura 4. Importância global (SHAP e PI).

Para avaliar a concordância entre os ranqueamentos globais gerados pelo SHAP e pelo PI, calculou-se o coeficiente de Spearman (ρ) para cada curso (Tabela 3). Observou-se correlação forte em ENF, BICT e CCO e correlação moderada — porém significativa — em ODO. A convergência entre os dois métodos, fundamentados em princípios distintos, indica estabilidade na identificação dos fatores mais relevantes e reforça a consistência interpretativa das explicações globais

Tabela 3. Correlação de Spearman entre SHAP e PI.

Curso	Spearman ρ	P-Value
ODO	0.67	0.0499
ENF	0.83	0.0053
BICT	0.88	0.0016
CCO	0.90	0.0011

Esses achados respondem diretamente à **RQ1**: *IECH, IEPL e MC* são, de forma consistente, os preditores mais influentes, enquanto atributos demográficos atuam de maneira secundária. Esse resultado é coerente com estudos que identificam indicadores de desempenho como os fatores mais estáveis na previsão de evasão [Barros et al. 2019] e [Chicon et al. 2025]. Em termos práticos, os achados sugerem que ações institucionais voltadas à permanência devem priorizar o acompanhamento da eficiência em aprovações e do ritmo de progressão, enquanto informações demográficas funcionam sobretudo como elementos de contexto para orientar intervenções mais amplas.

4.3. Explicabilidade local

Na análise local, as técnicas *SHAP*, *LIME* e *Anchors* foram utilizadas para interpretar decisões individuais dos modelos. A Figura 5 apresenta um estudante corretamente classificado como concluinte. O gráfico *waterfall* do *SHAP* e as barras do *LIME* destacam contribuições positivas de *IECH, IEPL e MC*, enquanto o *Anchors* resume esse padrão em regras de alta precisão que associam faixas elevadas desses indicadores a alta probabilidade de conclusão. As três técnicas convergem, portanto, para um perfil de forte desempenho acadêmico.

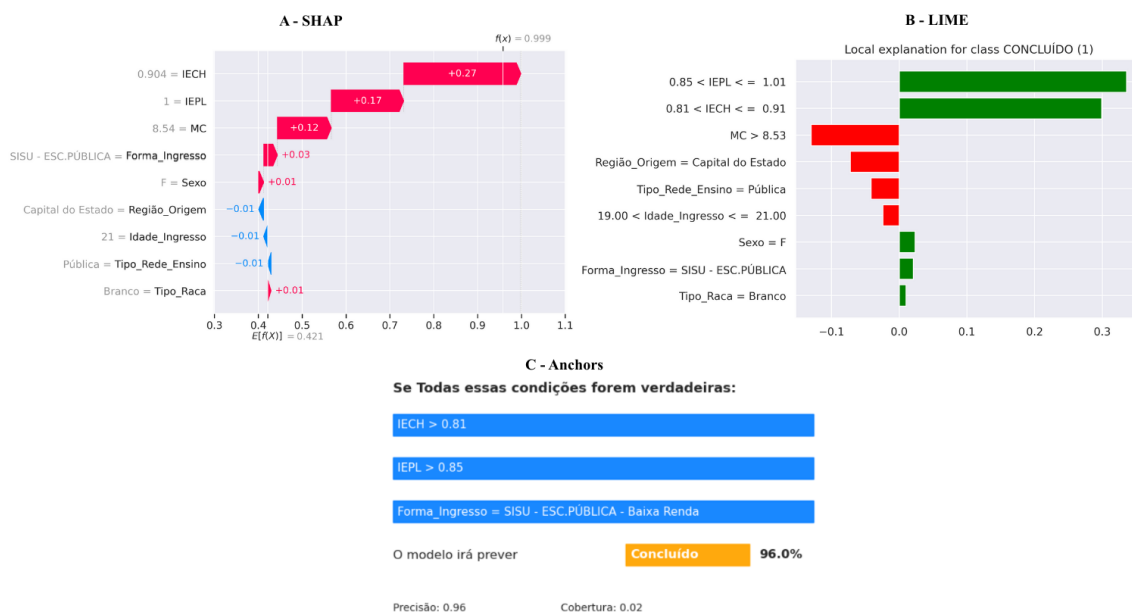


Figura 5. Comparação de explicações locais (SHAP, LIME e Anchors) para um caso corretamente previsto como concluído.

A Figura 6 mostra um caso corretamente previsto como desligamento. Nesse exemplo, valores reduzidos de *IECH* e *IEPL* aparecem como principais contribuintes negativos em *SHAP* e *LIME*, e o *Anchors* descreve esse cenário em regras que vinculam baixa eficiência e ritmo de progressão insuficiente a elevado risco de evasão. A recorrência desse padrão entre métodos distintos indica que os modelos capturam de forma consistente trajetórias acadêmicas bem-sucedidas e trajetórias em risco.

Respondendo à **RQ2**, as técnicas locais de XAI permitiram identificar pontos fortes e fragilidades no desempenho acadêmico em nível individual. Perfis com *IECH* e *IEPL* elevados associam-se a alta probabilidade de conclusão, enquanto valores baixos evidenciam zonas críticas que demandam intervenções precoces, como apoio acadêmico, tutoria ou ajuste de carga horária [Melo et al. 2022], [Figuroa-Cañas and Sancho-Vinuesa 2020] e [Chicon et al. 2025]. Embora devam ser interpretadas com cautela, devido à colinearidade entre indicadores e à ausência de garantias causais, essas análises ampliam a transparência do processo decisório e reforçam o potencial da XAI como suporte à gestão acadêmica.

4.4. Implicações Práticas da XAI

A interpretação das implicações deve considerar as limitações do conjunto de dados, que inclui apenas indicadores acadêmicos e demográficos. Fatores psicossociais, financeiros e motivacionais — amplamente reconhecidos como determinantes da evasão — não estavam disponíveis [Banaag et al. 2024]. Assim, as análises a seguir derivam das explicações geradas pelas técnicas de XAI sobre as decisões do modelo; as ações pedagógicas associadas baseiam-se em evidências da literatura e não são inferidas diretamente dos algoritmos.

As análises de explicabilidade revelaram três dimensões centrais nas decisões do modelo: eficiência na conversão da carga horária matriculada em aprovações (*IECH*), ritmo de progressão ao longo do curso (*IEPL*) e média de conclusão (*MC*). A análise combinada de *SHAP*, *PI*, *LIME* e *Anchors* esclareceu como cada variável contribuiu para

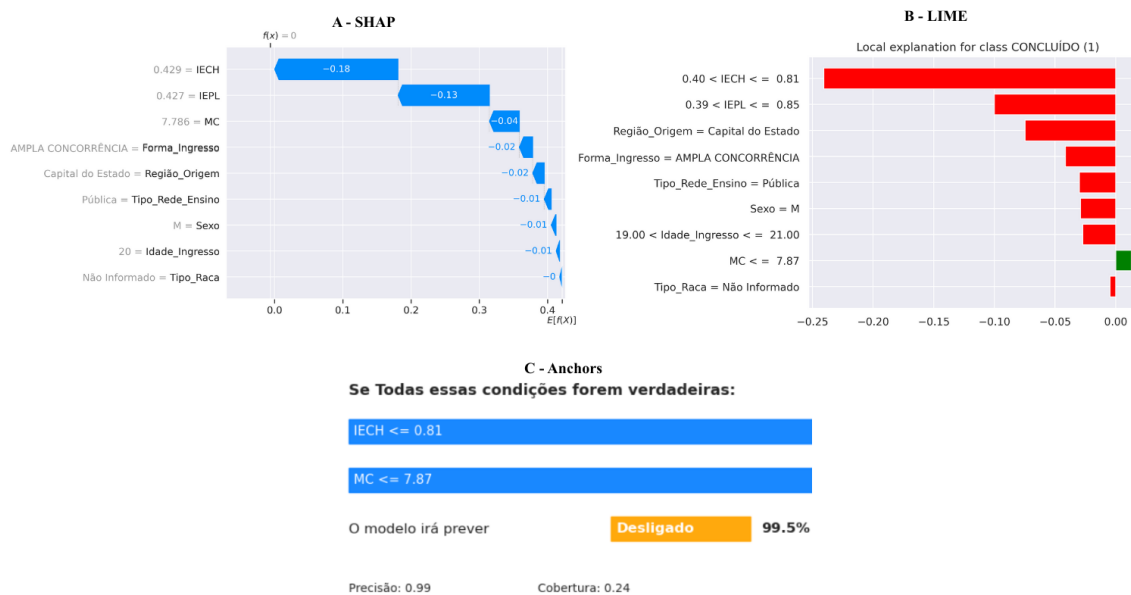


Figura 6. Comparação de explicações locais (SHAP, LIME e Anchors) para um caso corretamente previsto como desligado.

as predições de conclusão ou evasão, permitindo associar padrões explicativos a cenários específicos de desempenho.

- Faixas altas de *IECH* ($\geq 0,85$) e *IEPL* ($\geq 0,90$) apresentaram contribuições positivas consistentes para predições de conclusão. O modelo associa alta eficiência e progressão regular à maior probabilidade de permanência, perfil que a literatura relaciona a trajetórias consolidadas, para as quais ações preventivas e formativas, como feedback contínuo e estímulo ao engajamento, ajudam a manter o bom desempenho.
- Faixas intermediárias de *IECH* (entre 0,70 e 0,85) ou de *IEPL* (entre 0,75 e 0,90) produziram contribuições mais fracas, caracterizando um perfil de risco moderado, em que pequenas variações de rendimento podem alterar o desfecho. Evidências em [Saccaro et al. 2019] e [Melo et al. 2022] indicam a importância de monitoramento direcionado, ajustes de carga horária e participação em atividades de tutoria ou orientação acadêmica.
- Faixas baixas de *IECH* ($< 0,70$) e *IEPL* ($< 0,75$) geraram as contribuições negativas mais intensas nas predições de evasão. As análises locais de explicabilidade indicam que o modelo classifica esse perfil como de alto risco, refletindo baixo rendimento e progressão insuficiente. A literatura recomenda intervenções pedagógicas individualizadas, apoio psicopedagógico e reforço metodológico para esses casos — medidas eficazes para reverter trajetórias de evasão [Nagy and Molontay 2024], [Figueroa-Cañas and Sancho-Vinuesa 2020] e [Chicon et al. 2025].

Respondendo à **RQ3** as explicações fornecidas pelas técnicas de XAI convertem predições em evidências acionáveis para a gestão acadêmica. Ao indicar faixas de risco associadas à eficiência, progressão e rendimento, as análises permitem identificar grupos que demandam acompanhamento prioritário e orientar intervenções pedagógicas mais precisas, alinhadas às necessidades de cada curso.

Por fim, destaca-se a dimensão ética e pedagógica. As técnicas de explicabilidade não prescrevem ações, mas ampliam a compreensão das decisões do modelo, permitindo que gestores e docentes contextualizem casos e adotem intervenções proporcionais. A modelagem preditiva explicável, portanto, apoia sem substituir o julgamento humano, fornecendo base para políticas de monitoramento contínuo e para uma abordagem mais equitativa, transparente e responsável na retenção estudantil [Saarela and Podgorelec 2024] e [Altukhi and Pradhan 2025].

5. Conclusão

Os resultados deste estudo indicam que a combinação entre modelagem preditiva segmentada por curso e técnicas de XAI pode apoiar estratégias institucionais voltadas à prevenção da evasão acadêmica. De forma mais ampla, os achados mostram que métodos explicáveis ajudam a transformar previsões algorítmicas em informações compreensíveis para a tomada de decisão em contextos educacionais.

A principal contribuição metodológica consiste em um processo de interpretabilidade reutilizável, que integra preparação dos dados, modelagem segmentada por curso e técnicas globais e locais de explicabilidade, articuladas a um esquema de interpretação que vincula padrões identificados a possíveis intervenções pedagógicas. Esse processo pode ser adaptado por outras instituições, desde que sejam mapeados indicadores locais para equivalentes funcionais de eficiência e progressão e ajustados limiares de risco às políticas e estruturas curriculares de cada curso.

Do ponto de vista empírico, a análise indicou que a eficiência na conversão da carga horária em créditos aprovados (*IECH*), o ritmo de progressão (*IEPL*) e a média de conclusão (*MC*) são os principais determinantes das previsões, enquanto atributos demográficos exercem influência secundária. Esses achados sugerem que monitorar estudantes com progressão abaixo do esperado e fortalecer ações de nivelamento e tutoria podem contribuir para a redução do risco de evasão.

Em cursos com maiores taxas históricas de desligamento, como Ciência da Computação e BICT, os limiares de risco identificados pelos modelos permitem localizar grupos que demandam intervenção imediata. Em programas com menor evasão, como Enfermagem e Odontologia, esses indicadores permitem detectar precocemente mudanças na trajetória acadêmica. Assim, o processo proposto não apenas explica padrões de evasão, mas também oferece base objetiva para qualificar práticas pedagógicas e orientar decisões curriculares voltadas à redução da evasão e ao apoio à conclusão.

Este estudo apresenta algumas limitações. A ausência de variáveis psicossociais, socioeconômicas e de engajamento estudantil restringe a análise às dimensões acadêmicas, e o uso de dados de uma única instituição pública pode limitar a generalização dos achados. Além disso, não foram exploradas estratégias avançadas de ajuste de hiperparâmetros, o que pode ter reduzido o desempenho de alguns algoritmos. As explicações produzidas pelos métodos de XAI descrevem associações estatísticas aprendidas pelos modelos e não devem ser interpretadas como evidências causais da evasão. Estudos futuros podem ampliar o escopo amostral, incorporar novas abordagens de modelagem e investigar estratégias de validação causal.

Agradecimentos

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão - FAPEMA e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Uso de Inteligência Artificial

Este trabalho fez uso de ferramenta de Inteligência Artificial generativa (ChatGPT) como apoio técnico à organização das informações. Em particular, a ferramenta foi empregada para apoiar a implementação e a padronização de códigos utilizados na geração de gráficos explicativos dos métodos de XAI (SHAP, LIME e Anchors), bem como na organização comparativa das visualizações resultantes.

Destaca-se que todas as contribuições da ferramenta de IA foram supervisionadas, avaliadas criticamente e ajustadas pelos autores, não tendo a ferramenta participado da definição metodológica, da análise dos resultados ou da interpretação científica dos achados, os quais são de responsabilidade exclusiva dos autores.

Referências

- Altukhi, Z. M. and Pradhan, S. (2025). Systematic literature review: Explainable AI definitions and challenges in education. *arXiv*.
- Banaag, R., Sumodevilla, J. L., and Potane, J. (2024). Factors affecting student dropout behavior: A systematic review. *International Journal of Educational Management and Innovation*, 5(1):53–70.
- Barbosa, T., Freitas, N., Cavalcanti, L., Batista, M. C. M., Gouveia, R., and Alves, G. (2024). Construção dinâmica de modelos de learning analytics explicáveis e justos aplicados ao acompanhamento de estudantes de graduação. In *Proc. SBIE*.
- Barros, T. M., Pacheco, R. A. C., de Oliveira, A. C., Correa, D. D., and dos Santos, J. A. (2019). Determination of dropout student profile based on correspondence analysis technique. *IEEE Latin America Transactions*, 17(9):1517–1523.
- Chicon, P. M. M., Paschoal, L. N., Sawicki, S., Roos-Frantz, F., and Frantz, R. Z. (2025). A predictive model for the early identification of student dropout using data classification, clustering, and association methods. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 20:12–21.
- da Silva, F. C., Feitosa, R. M., Batista, L. A., and Santana, A. M. (2024). Análise comparativa de métodos de explicabilidade da inteligência artificial no cenário educacional: Um estudo de caso sobre evasão. In *Proc. SBIE*, pages 2968–2977.
- da Silva, W. S., de Oliveira, V. T., Araújo, S. S., Vieira, D., and de Castro, M. F. (2022). Explainability e auditability: Interpretando e validando modelos de machine learning. In *Proc. SBSI*, pages 456–467.
- da Silva Dantas, J. C., da Silva Aranha, E. H., and da Silva, T. R. (2024). Predição de alunos em risco de reprovação: Uma comparação do impacto de diferentes técnicas de amostragem. *RENOTE*, 22(3):230–239.

- Farrow, R. (2023). The possibilities and limits of explicable artificial intelligence (XAI) in education: A socio-technical perspective. *Learning, Media and Technology*, 48(2):266–279.
- Figueroa-Cañas, J. and Sancho-Vinuesa, T. (2020). Early prediction of dropout and final exam performance in an online statistics course. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 15(2):86–94.
- Fiok, K., Farahani, F. V., Karwowski, W., and Ahram, T. (2022). Explainable artificial intelligence for education and training. *Journal of Defense Modeling and Simulation*, 19(2):133–144.
- Géron, A. (2021). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O’Reilly Media, 2nd edition.
- Harrison, M. (2020). *Machine learning pocket reference: Working with structured data in Python*. O’Reilly Media.
- Heredia, D., Amaya, Y., and Barrientos, E. (2015). Student dropout predictive model using data mining techniques. *IEEE Latin America Transactions*, 13(9):3127–3134.
- Hernández-Gonzalez, A. G., Meléndez-Armenta, R. A., Morales-Rosales, L. A., Barrientos, A. G., Tecpanecatl-Xihuitl, J. L., and Algreto, I. (2016). Comparative study of algorithms to predict student desertion at ITSM-Mexico. *IEEE Latin America Transactions*, 14(11):4573–4578.
- Klockmann, V., von Schenk, A., and Villeval, M. C. (2022). Artificial intelligence, ethics, and intergenerational responsibility. *Journal of Economic Behavior and Organization*, 203:284–317.
- Kocsis, Á. and Molnár, G. (2025). Factors influencing academic performance and dropout rates in higher education. *Oxford Review of Education*, 51(3):414–432.
- Lundberg, S. M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proc. NeurIPS*.
- Marques, J., Gonçalves, C., Vieira, P. A., Borges, A., Dias, V., Santos, W., and Silva, R. (2023). Explorando a explicabilidade da inteligência artificial: Técnicas para compreender e interpretar modelos de aprendizado de máquina. In *Título da coletânea*, pages 25–43. SBC.
- Melo, E., Silva, I., Costa, D. G., Viegas, C. M. D., and Barros, T. M. (2022). On the use of explainable artificial intelligence to evaluate school dropout. *Education Sciences*, 12(12):845.
- Molnar, C. (2025). *Interpretable machine learning: A guide for making black box models explainable*. 3rd edition.
- Nagy, M. and Molontay, R. (2024). Interpretable dropout prediction: Towards XAI-based personalized intervention. *International Journal of Artificial Intelligence in Education*, 34(2):274–300.
- Neugebauer, M., Becker, M., Bihler, L.-M., and Wagner, J. (2025). Well-being scarring effects of college non-completion. *Social Science Research*, 127:103138.

- Niyogisubizo, J., Liao, L., Nziyumva, E., Murwanashyaka, E., and Nshimyumukiza, P. C. (2022). Predicting student dropout using a two-layer ensemble machine learning approach: A novel stacked generalization. *Computers and Education: Artificial Intelligence*, 3:100066.
- OECD (2023). *Education at a glance 2023: OECD indicators*. OECD Publishing.
- Pérez, B., Castellanos, C., and Correal, D. (2018). Predicting student drop-out rates using data mining techniques: A case study. In *Proc. IEEE Colombian Conf. Applications in Computational Intelligence*, pages 111–125.
- Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). Why should i trust you? explaining the predictions of any classifier. In *Proc. KDD*, pages 1135–1144.
- Ribeiro, M. T., Singh, S., and Guestrin, C. (2018). Anchors: High-precision model-agnostic explanations. In *Proc. AAAI*, volume 32.
- Rodrigues, H. S., Martins, A. C., Silva, J. P., and Rocha, M. L. (2024). Investigating AI algorithms to predict college students' academic performance: A systematic mapping study. In *Proc. ICAART*.
- Saarela, M. and Podgorelec, V. (2024). Recent applications of explainable AI (XAI): A systematic literature review. *Applied Sciences*, 14(19):8884.
- Saccaro, A., França, M. T. A., and Jacinto, P. d. A. (2019). Fatores associados à evasão no ensino superior brasileiro. *Estudos Econômicos*, 49(2):337–373.
- Saito, T. and Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3):e0118432.
- Saqr, M. and López-Pernas, S. (2024). Why explainable AI may not be enough: Predictions and mispredictions in decision making in education. *Smart Learning Environments*, 11(1):52.
- Semesp, I. (2025). Mapa do ensino superior no brasil 2025. Relatório técnico.
- Tete, M. F., Sousa, M. M., de Santana, T. S., and da Silva, S. F. (2022). Predictive models for higher education dropout: A systematic literature review. *Education Policy Analysis Archives*, 30(61).
- Ulloa-Cázares, R. L., García-Díaz, N., and Soriano-Equigua, L. (2021). Multi-layer adaptive fuzzy inference system for predicting student performance in online higher education. *IEEE Latin America Transactions*, 19(1):98–106.