

Modelagem Preditiva da Permanência Acadêmica de Mães Universitárias com Aprendizado de Máquina

Ana G. B. de Lima¹, Andressa B. Ferreira¹, Gabriel J. S. Duarte¹

¹ Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE) – Campus Aracati
Curso de Bacharelado em Ciência da Computação
Aracati – CE – Brasil

`ana.gabrielle.brandao62@aluno.ifce.edu.br`

`andressa.ferreira@ifce.edu.br, gabriel.duarte@ifce.edu.br`

1. Contexto e Motivação

A evasão e a permanência no ensino superior constituem fenômenos complexos e multifatoriais, amplamente investigados nas áreas da Educação e da Computação, especialmente com o avanço das técnicas de Mineração de Dados Educacionais e Aprendizado de Máquina [1, 2]. Essas abordagens permitem identificar padrões e fatores associados ao desempenho acadêmico e à trajetória dos estudantes, contribuindo para o desenvolvimento de estratégias institucionais de apoio à permanência.

Entretanto, determinados grupos sociais permanecem sub-representados em análises quantitativas, como mulheres que conciliam a maternidade com a formação universitária. Estudos apontam que a maternidade impõe desafios adicionais relacionados à organização do tempo, às condições socioeconômicas, à fragilidade das redes de apoio e às responsabilidades familiares, fatores que impactam diretamente a permanência acadêmica [3, 8]. No contexto brasileiro, políticas de assistência estudantil buscam mitigar parte dessas desigualdades, embora ainda existam lacunas na identificação precoce de fatores associados ao risco de evasão [9, 10].

2. Problema de Pesquisa

Apesar do crescimento de estudos voltados à predição da evasão acadêmica, observa-se uma escassez de pesquisas que considerem, de forma específica, a realidade de estudantes que são mães, integrando variáveis acadêmicas, financeiras, emocionais, de apoio social e de responsabilidades familiares. Além disso, muitos trabalhos priorizam modelos preditivos complexos, com baixa interpretabilidade, o que pode limitar sua aplicação em contextos educacionais sensíveis do ponto de vista social e ético [7, 11].

Diante desse cenário, surge a necessidade de investigar como técnicas de Aprendizado de Máquina interpretáveis podem apoiar a compreensão dos fatores associados à permanência ou à desistência acadêmica de mães universitárias, contribuindo para análises mais transparentes e socialmente responsáveis.

3. Objetivos

O objetivo geral deste trabalho é propor uma abordagem de modelagem preditiva baseada em Aprendizado de Máquina para analisar a permanência acadêmica de mães universitárias. Como objetivos específicos, busca-se identificar variáveis relevantes associadas à trajetória acadêmica desse público, estruturar um conjunto de dados a partir de

informações autorrelatadas e aplicar modelos supervisionados de classificação binária que priorizem a interpretabilidade dos resultados.

4. Metodologia Proposta

A pesquisa adota uma abordagem quantitativa, com coleta de dados via questionário estruturado aplicado a mulheres que sejam mães e estejam matriculadas, tenham evadido ou já concluído o ensino superior. Considera-se como “mãe universitária” toda mulher que vivenciou a maternidade durante o período de sua formação acadêmica.

O instrumento contempla dimensões acadêmicas, financeiras, emocionais, de apoio social e de responsabilidades familiares. O recrutamento ocorrerá via redes acadêmicas e sociais, buscando diversidade de cursos e instituições, com tamanho amostral mínimo estimado em 100 respondentes.

Os dados serão submetidos a pré-processamento (limpeza, codificação e análise exploratória). A variável-alvo será operacionalizada como classificação binária (permanência ou desistência): participantes que concluíram o curso serão classificadas como permanência; casos intermediários serão tratados com critérios específicos para garantir consistência na rotulagem.

A modelagem preditiva utilizará Regressão Logística [4], Random Forest [5] e Support Vector Machine [6], selecionados pelo equilíbrio entre desempenho preditivo e simplicidade estrutural. A explicabilidade é aspecto central: serão adotadas técnicas como LIME [12] e SHAP [13], que permitem analisar contribuições individuais e globais das variáveis, favorecendo transparência e interpretabilidade [7]. A pesquisa será submetida ao Comitê de Ética em Pesquisa (CEP) competente, garantindo anonimato, consentimento informado e uso responsável dos dados.

5. Contribuições Esperadas

Espera-se que este estudo contribua para a compreensão dos fatores associados à permanência acadêmica de mães universitárias, oferecendo subsídios para a formulação de ações institucionais mais direcionadas, como políticas de assistência estudantil, estratégias de apoio à conciliação entre maternidade e vida acadêmica e mecanismos de identificação precoce de risco de evasão.

Além disso, o uso de modelos interpretáveis busca apoiar decisões baseadas em evidências, promovendo maior transparência e responsabilidade social na aplicação de técnicas de Aprendizado de Máquina no ensino superior.

Como complemento à proposta apresentada, foi elaborado um vídeo explicativo no qual são discutidos o contexto, a motivação pessoal da pesquisa e os principais elementos metodológicos do estudo¹.

¹<https://www.youtube.com/watch?v=qdExThoShf8>

Referências

- [1] ROMERO, C.; VENTURA, S. Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, v. 40, n. 6, p. 601–618, 2010.
- [2] BAKER, R.; INVENTADO, P. S. Educational data mining and learning analytics. In: SAWYER, R. K. (org.). *The Cambridge Handbook of the Learning Sciences*. 2. ed. Cambridge: Cambridge University Press, 2014. p. 589–607.
- [3] LEITE, A. C. F.; ALVES, F. C. Trabalho, maternidade e permanência no ensino superior. *Ensino em Perspectivas*, v. 3, n. 1, p. 1–15, 2022.
- [4] HOSMER, D. W.; LEMESHOW, S. *Applied Logistic Regression*. 2. ed. New York: Wiley, 2000.
- [5] BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.
- [6] CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, v. 20, n. 3, p. 273–297, 1995.
- [7] LINARDATOS, P. et al. Explainable AI: a review of machine learning interpretability methods. *Entropy*, v. 23, n. 1, 2020.
- [8] STANISCUASKI, F. et al. Bias against parents in science hits women harder. *Humanities & Social Sciences Communications*, v. 10, n. 201, 2023.
- [9] BRASIL. Ministério da Educação. *Plano Nacional de Assistência Estudantil (PNAES)*. Brasília: MEC, 2010.
- [10] INEP. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo da Educação Superior 2022: resumo técnico*. Brasília: INEP, 2023.
- [11] STRÜMKE, I.; SLAVKOVIK, M. Explainability for identification of vulnerable groups in machine learning models. *arXiv preprint arXiv:2203.00317*, 2022.
- [12] RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. “Why should I trust you?”: Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2016. p. 1135–1144.
- [13] LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, v. 30, 2017.