

Análise da trajetória acadêmica estudantil utilizando mineração de processos com ênfase na evasão do curso de Ciência da Computação da UFC em Quixadá

Antonio Gildo de Freitas Filho¹, Paulo T. Guerra¹, Davi R. Vasconcelos¹

¹Universidade Federal do Ceara (UFC)

Av. José de Freitas Queiroz, 5003 – Cedro – 63902-580 – Quixadá — CE — Brazil

`gildofreitas@alu.ufc.br, {paulodetarso,daviromero}@ufc.br`

Abstract. *This study analyses a formal model of the Computer Science curriculum at UFC – Quixadá Campus using Petri nets. Based on students’ academic records, an event log was constructed to represent their trajectories over the semesters. Using the token replay technique, metrics were extracted to compare graduate and dropout students. The results reveal distinct patterns of adherence to the curricular model and allow the identification of bottlenecks and critical points associated with student attrition.*

Resumo. *Este trabalho analisa uma modelagem da estrutura curricular do curso de Ciência da Computação da UFC – Campus de Quixadá, por meio de redes de Petri. Com base nos históricos acadêmicos dos discentes, foi construído um log de eventos representando suas trajetórias ao longo dos semestres. Utilizando a técnica de token replay, foram extraídas métricas que possibilitaram comparar estudantes formados e evadidos. Os resultados evidenciam padrões distintos de aderência ao modelo curricular e permitem identificar gargalos e pontos críticos associados à evasão.*

1. Introdução

A evasão de estudantes no ensino superior permanece como um desafio crítico enfrentado por instituições educacionais, especialmente no contexto brasileiro. O abandono precoce de cursos de graduação compromete não apenas os projetos de vida dos discentes, mas também a eficiência institucional e o financiamento público. Nesse contexto, compreender os caminhos percorridos pelos estudantes ao longo de sua trajetória acadêmica é essencial para identificar gargalos curriculares e subsidiar estratégias de intervenção mais eficazes.

Nos últimos anos, o avanço de tecnologias voltadas à análise de dados educacionais tem permitido novas abordagens para investigar esse problema. A Mineração de Processos (*Process Mining*) [Van der Aalst 2016], ao integrar conceitos da Ciência de Dados e modelagem de processos, possibilita analisar como fluxos reais de eventos se relacionam com modelos formais da trajetória acadêmica ideal definida pela estrutura curricular. Nesse cenário, as redes de Petri se destacam por sua capacidade de representar formalmente processos com estados e transições, sendo adequadas para modelar estruturas curriculares no ensino superior.

Este trabalho propõe uma abordagem baseada na técnica de *token replay* aplicada à formalização da estrutura curricular de um curso de Ciência da Computação por meio

de redes de Petri. A técnica permite verificar, a partir de *logs* de eventos compostos pelos históricos acadêmicos dos estudantes, o grau de conformidade entre os percursos realizados e o fluxo curricular esperado. Além disso, métricas como *tokens* consumidos, ausentes, produzidos e remanescentes, bem como o índice de *fitness*, possibilitam identificar desvios, gargalos e padrões associados à evasão acadêmica.

O restante deste artigo está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 descreve os fundamentos teóricos utilizados; a Seção 4 detalha a metodologia adotada, incluindo a modelagem do currículo e a aplicação do *token replay*; a Seção 5 apresenta os resultados e análises realizadas; por fim, a Seção 6 traz as considerações finais, limitações do estudo e sugestões de trabalhos futuros.

2. Trabalhos Relacionados

Os trabalhos relacionados foram selecionados a partir de buscas em bases da área de Computação e Educação, utilizando termos associados à evasão estudantil, mineração de processos e trajetórias acadêmicas.

Em [do Carmo et al. 2022], foram analisadas trajetórias acadêmicas de estudantes de Ciência da Computação da UDESC por meio de padrões sequenciais e análise de grafos, identificando elevada evasão nos semestres iniciais após reprovações em disciplinas introdutórias. Já [Unger et al. 2022] aplicou mineração de processos em ambientes virtuais de aprendizagem para identificar padrões de interação associados ao desempenho acadêmico. Por sua vez, [Kantorski et al. 2023] utilizou algoritmos de aprendizado de máquina para prever risco de evasão com base em dados históricos educacionais.

Embora essas abordagens contribuam para a compreensão da evasão, grande parte dos estudos concentra-se em modelos preditivos ou análises comportamentais, com menor ênfase na representação explícita do fluxo curricular. Diferentemente desses trabalhos, este estudo utiliza redes de Petri e a técnica de *token replay* para modelar trajetórias acadêmicas de forma sequencial, permitindo identificar desvios, gargalos curriculares e padrões associados à evasão a partir da aderência ao fluxo curricular esperado.

O presente trabalho constitui um recorte de uma pesquisa de mestrado mais ampla [de Freitas Filho 2025], na qual foi desenvolvida uma modelagem formal do fluxo curricular do curso de Ciência da Computação da Universidade Federal do Ceará (UFC), campus Quixadá, utilizando dados reais de históricos acadêmicos. Neste recorte, enfatiza-se a análise das métricas de aderência obtidas por meio do *token replay*, buscando compreender padrões de retenção e evasão no percurso acadêmico discente.

3. Fundamentação teórica

Esta seção apresenta os principais fundamentos teóricos relacionados à evasão no ensino superior, à mineração de processos e à técnica de *token replay*.

3.1. Evasão no ensino superior

Segundo [INEP 2017], a evasão em cursos de graduação caracteriza-se pela desistência do estudante antes da conclusão do curso. Esse fenômeno produz impactos sociais, pessoais e institucionais relevantes [Souza et al. 2015, Saraiva et al. 2021, Kantorski et al. 2023].

O indicador de evasão é acompanhado pelo Censo da Educação Superior, realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP),

que monitora o fluxo discente ao longo de dez anos após o ingresso. Em 2024, a taxa de evasão no ensino superior brasileiro atingiu 59% para a *coorte* de 2014–2023. Essa *coorte* representa o acompanhamento longitudinal dos estudantes considerando três condições: permanência, conclusão ou desistência.

No curso de Ciência da Computação, a *coorte* de 2014 apresentou taxa nacional de evasão de 71,8%, indicando que, para cada estudante concluinte, quase três abandonaram o curso. Na Universidade Federal do Ceará (UFC), campus Quixadá, a taxa atingiu 76,2%.

Diversos estudos investigam fatores associados à evasão, incluindo baixo rendimento acadêmico [Érica Carmo et al. 2022, Viana et al. 2022], aspectos socioemocionais [Saraiva et al. 2021], mudança de interesses [Morais et al. 2021], distância geográfica da instituição [Colpo et al. 2021], problemas institucionais [Carvalho et al. 2023] e características da própria estrutura curricular [dos Santos Oliveira and de Medeiros 2024, Jesus et al. 2014].

Neste trabalho, considera-se que a estrutura curricular pode ser modelada como um processo, permitindo aplicar técnicas de análise de conformidade para identificar gargalos e desvios ao longo da trajetória acadêmica dos estudantes.

3.2. Mineração de processos

A análise de trajetórias acadêmicas a partir de dados educacionais demanda modelos capazes de representar formalmente o fluxo esperado dos estudantes ao longo do curso. Nesse contexto, as redes de Petri destacam-se por representar processos baseados em estados e transições, sendo adequadas para modelar estruturas curriculares.

Segundo [Van der Aalst 2016], uma rede de Petri é um grafo bipartido composto por lugares, transições e arcos direcionados que definem o fluxo do processo. Formalmente, uma rede de Petri é definida como uma tripla $N = (P, T, F)$, em que:

- **P** representa um conjunto finito de lugares;
- **T** representa um conjunto finito de transições, tal que $P \cap T = \emptyset$;
- **F** representa o conjunto de arcos direcionados, de modo que $F \subseteq (P \times T) \cup (T \times P)$.

A dinâmica da rede é representada pela movimentação de *tokens* entre lugares e transições, permitindo modelar a execução do processo e a evolução dos estados ao longo do fluxo curricular [Van der Aalst 2016].

Na mineração de processos, técnicas de análise de conformidade permitem comparar registros reais de execução com um modelo formal do processo. Entre essas técnicas, o *token replay* destaca-se por possibilitar a verificação da aderência entre trajetórias observadas e o fluxo esperado, permitindo identificar desvios e gargalos ao longo do percurso acadêmico.

A partir de um *log* de eventos estruturado e de um modelo de processo definido, o *token replay* executa cada *trace* no modelo e calcula métricas associadas ao grau de conformidade entre o comportamento observado e o comportamento esperado. Ao final da execução, é inferido um valor de *fitness* para cada *trace*, conforme definido pela Equação 1.

$$fitness(\sigma, N) = 0,5 \left(1 - \frac{m}{c}\right) + 0,5 \left(1 - \frac{r}{p}\right) \quad (1)$$

A métrica de *fitness* considera a relação entre *tokens* consumidos, produzidos, ausentes e remanescentes durante a execução do *token replay*. Valores mais altos indicam maior aderência entre o percurso acadêmico observado e o fluxo curricular modelado, enquanto desvios sinalizam possíveis gargalos ou inconsistências nas trajetórias analisadas.

4. Metodologia

O experimento foi implementado no ambiente *Google Colaboratory*¹, utilizando a linguagem *Python*² para pré-processamento dos dados e análise dos resultados. O *software Snoopy*³ foi utilizado na modelagem da rede de Petri, enquanto a biblioteca *PM4PY*⁴ foi empregada na execução do *token replay*. O processo metodológico foi organizado em cinco etapas: (i) coleta e pré-processamento dos dados; (ii) estruturação do *log* de eventos; (iii) modelagem da estrutura curricular; (iv) aplicação do *token replay*; e (v) análise dos resultados.

A etapa (i) coleta de dados e pré-processamento foi realizada a partir da extração do histórico acadêmico dos discentes no Sistema Integrado de Gestão de Atividades Acadêmicas (SIGAA), disponibilizado pela Pró-Reitoria de Graduação (PROGRAD) do curso de Bacharelado em Ciência da Computação da Universidade Federal do Ceará (UFC), campus Quixadá. O estudo utilizou exclusivamente dados acadêmicos anonimizados e agregados, sem identificação dos discentes e sem intervenção direta com seres humanos. O conjunto analisado totalizou 110 estudantes. As atividades de pré-processamento incluíram:

1. definição das variáveis de interesse: *ano*, *periodo*, *id_discente*, *codigo* e *descricao*;
2. remoção de registros duplicados gerados por turmas compartilhadas entre docentes;
3. exclusão de disciplinas optativas;
4. conversão das variáveis *ano* e *periodo* na variável *Semestre Letivo*;
5. padronização dos *status* finais das disciplinas em três categorias: “MATRICULADO”, “APROVADO” e “NAOAPROVADO”;
6. padronização dos códigos de disciplinas equivalentes.

A Tabela 1 apresenta um exemplo do conjunto de dados após a etapa de pré-processamento.

Na etapa (ii), o *log* de eventos foi estruturado considerando os atributos básicos de mineração de processos: *Case Id* (*id_discente*), *Activity* (disciplinas, semestres e status acadêmicos) e *Timestamp*. As variáveis *codigo* e *descricao* foram concatenadas, além da criação das atividades *semestre* e *formado*, responsáveis por representar a progressão temporal e a conclusão do curso. A Tabela 2 apresenta um exemplo do *log* de eventos estruturado para um discente genérico.

¹<https://colab.google/>

²<https://www.python.org/>

³<https://www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Snoopy>

⁴<https://pypi.org/project/pm4py/>

Tabela 1. Exemplo do conjunto de dados após pré-processamento

codigo	id_discente	descricao	Semestre_letivo
QXD0005	999003	NAOAPROVADO	2014-02-01
QXD0103	999141	NAOAPROVADO	2014-02-01
QXD0001	999003	NAOAPROVADO	2014-02-01
QXD0108	999448	APROVADO	2014-02-01
QXD0056	999208	NAOAPROVADO	2014-02-01
		⋮	

Tabela 2. Log de eventos do discente 1003

timestamp	id_discente	codigo
2015-01-01	1003	semestre01
2015-02-01	1003	QXD0001
2015-02-01	1003	QXD0105
2015-02-01	1003	QXD0056
2015-02-01	1003	QXD0001_NAOAPROVADO
2015-02-01	1003	QXD0105_NAOAPROVADO
2015-02-01	1003	QXD0056_NAOAPROVADO
2015-07-01	1003	semestre02
2015-02-01	1003	QXD0001_APROVADO
2015-02-01	1003	QXD0105_APROVADO
2015-02-01	1003	QXD0056_APROVADO
2015-07-01	1003	semestre03
...
2019-06-03	1003	formado

Na etapa (iii), foi construída uma rede de Petri representando a estrutura curricular do curso. A Figura 1 apresenta uma versão reduzida do modelo, contendo lugar inicial, lugar final e transições associadas aos semestres, disciplinas e respectivos *status* acadêmicos.

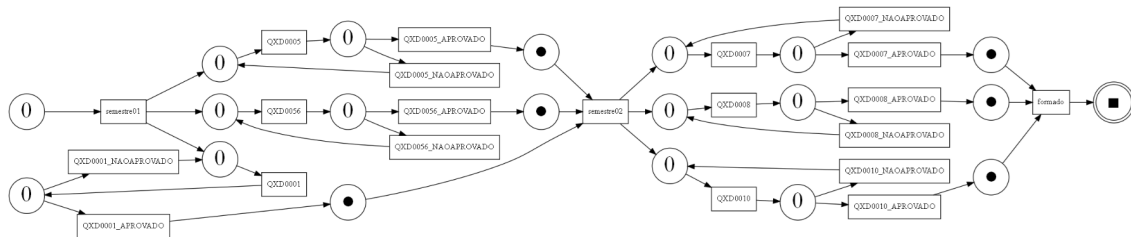


Figura 1. Modelo reduzido da estrutura curricular do curso de Ciência da Computação

A progressão dos estudantes ocorre conforme a execução das transições registradas no *log* de eventos. Situações de reprovação ou desvios do fluxo curricular esperado podem gerar *tokens* ausentes ou remanescentes, refletindo inconsistências entre a trajetória observada e o modelo curricular definido. Dessa forma, a aplicação do *token replay* permite identificar padrões de retenção, gargalos curriculares e trajetórias associadas à evasão acadêmica.

5. Experimentos

A aplicação da técnica de *token replay* sobre a rede de Petri modelada a partir da estrutura curricular do curso de Ciência da Computação permitiu analisar a conformidade entre as trajetórias acadêmicas observadas e o fluxo curricular previsto.

Em um cenário ideal, no qual o discente percorre integralmente a estrutura curricular prevista e obtém aprovação em todas as disciplinas, o modelo apresenta aderência máxima ao fluxo curricular, com valor de *fitness* igual a 1. Nessa situação, seriam produzidos e consumidos 123 *tokens*, sem ocorrência de *tokens* ausentes ou remanescentes. A partir dessa referência, as análises foram organizadas em dois grupos: discentes evadidos e formados.

Os discentes formados apresentam médias de produção (136,5) e consumo (142,2) de *tokens* superiores ao valor ideal, indicando ocorrências de reprovação ao longo da trajetória acadêmica. Entre os evadidos, os valores médios de produção (95,3) e consumo (99,5) foram inferiores aos observados entre os formados, sugerindo interrupção antecipada do percurso curricular. A Figura 2 apresenta as distribuições dessas métricas.

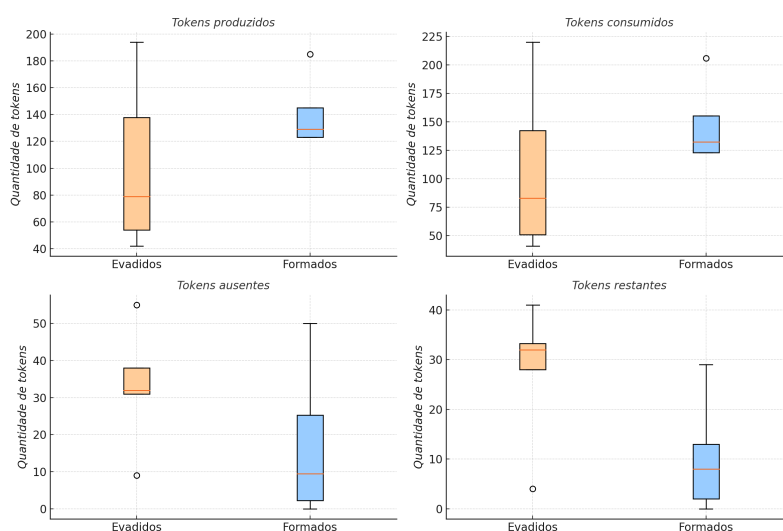


Figura 2. Estatísticas dos discentes evadidos e formados

Observa-se ainda que os discentes formados mantêm trajetórias relativamente estáveis, com menor dispersão das métricas, enquanto os evadidos apresentam maior variabilidade, refletindo percursos acadêmicos interrompidos em diferentes momentos do curso.

No que se refere aos *tokens* ausentes e remanescentes, os discentes formados apresentam médias reduzidas — 14,87 e 9,21, respectivamente — associadas a ajustes pontuais na trajetória acadêmica. Em contraste, os discentes evadidos apresentam médias de *tokens* ausentes (33,85) e remanescentes (29,59) significativamente superiores, indicando recorrência de reprovações e baixa aderência ao fluxo curricular esperado.

A Figura 3 apresenta a distribuição dos valores de *trace fitness*. Os discentes formados apresentam alta conformidade com o modelo formal, concentrando valores próximos de 1. Já entre os evadidos, observa-se maior dispersão, associada a múltiplas reprovações e interrupções no percurso acadêmico.

Os menores valores de *fitness* indicam evasões nos semestres iniciais, enquanto valores intermediários representam estudantes que avançaram parcialmente na matriz curricular, mas permaneceram acumulando retenções em disciplinas de períodos anteriores. Esses resultados sugerem que a evasão não ocorre apenas de forma abrupta nos semestres

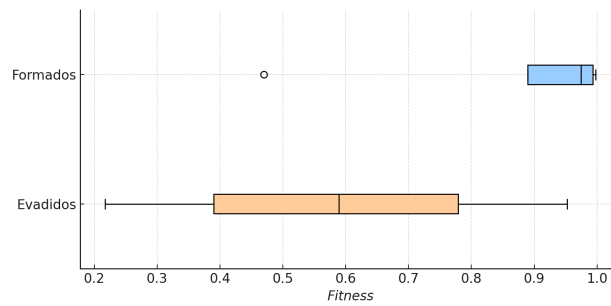


Figura 3. Estatísticas do *trace fitness* dos discentes evadidos e formados

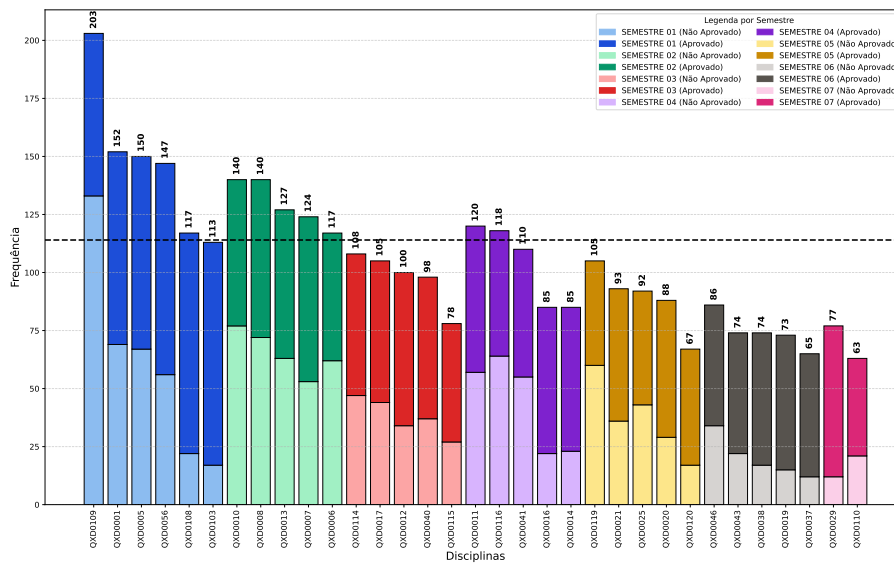


Figura 4. Frequência das transições por disciplinas

iniciais, mas também como consequência de sucessivas retenções ao longo da trajetória acadêmica. Nesse contexto, disciplinas com elevados índices de reprovação podem atuar como pontos críticos da estrutura curricular, indicando a necessidade de ações institucionais de acompanhamento acadêmico, monitorias e revisão de estratégias pedagógicas.

A Figura 4 apresenta a frequência de ativação das transições associadas às disciplinas durante a execução do *token replay*. Em um cenário ideal, cada disciplina deveria apresentar exatamente 110 ativações, correspondente ao total de discentes analisados. Observa-se maior frequência de ativação nas disciplinas dos primeiros semestres, indicando recorrência de reprovações e necessidade de recorrer componentes curriculares introdutórios.

Destacam-se as disciplinas *QXD0109* (Pré-Cálculo), *QXD0010* (Estrutura de Dados) e *QXD0008* (Matemática Discreta) como principais gargalos curriculares nos semestres iniciais. No 4º semestre, disciplinas como *QXD0011* (Fundamentos de Banco de Dados) e *QXD0116* (Álgebra Linear) apresentam frequência superior ao esperado, indicando reorganizações na trajetória acadêmica. Esse comportamento evidencia maior flexibilidade em disciplinas sem pré-requisitos rígidos, enquanto componentes associados a cadeias de dependência mais restritivas tendem a concentrar retenções sucessivas e maior risco de evasão.

A Figura 5 apresenta a distribuição dos *tokens* remanescentes ao longo da estrutura curricular. Observa-se aumento progressivo desses valores nos semestres finais do curso, indicando acúmulo de retenções e interrupções do percurso acadêmico.

Ao analisar algumas ramificações definidas pelas relações de pré-requisito, verifica-se aumento acumulativo de *tokens* remanescentes ao longo da progressão curricular. Um exemplo ocorre no caminho $QXD0001 \rightarrow QXD0010 \rightarrow QXD0041$, em que os *tokens* remanescentes aumentam de 31 para 62. Padrões semelhantes são observados em outras ramificações curriculares, como: $QXD0001 \rightarrow QXD0007 \rightarrow QXD0020$, $QXD0019$ ou $QXD0014$; $QXD0005 \rightarrow QXD0013 \rightarrow QXD0043$; $QXD0056 \rightarrow QXD0008 \rightarrow QXD0041$; $QXD0056 \rightarrow QXD0008 \rightarrow QXD0040 \rightarrow QXD0025$.

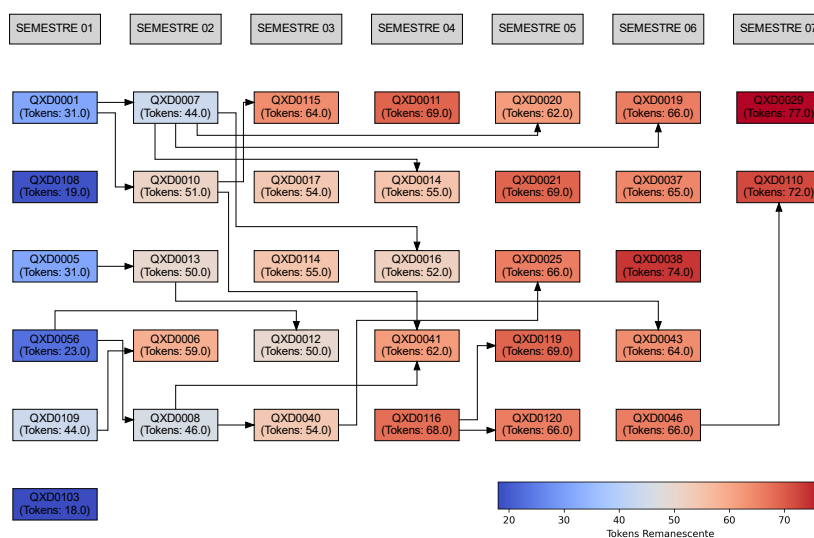


Figura 5. Distribuição de *tokens* remanescentes por disciplinas

Esses padrões indicam que os efeitos da retenção tendem a se propagar ao longo da matriz curricular, comprometendo disciplinas subsequentes e ampliando o risco de evasão. Assim, a identificação dessas ramificações pode auxiliar gestores acadêmicos na priorização de ações preventivas em componentes curriculares considerados críticos.

Como limitação, o estudo considera exclusivamente dados acadêmicos estruturados, sem incorporar fatores socioeconômicos, emocionais ou comportamentais que também influenciam a evasão. Além disso, a análise foi realizada em um único curso de graduação, embora a abordagem possa ser adaptada para outras estruturas curriculares semelhantes.

Em síntese, a técnica de *token replay* mostrou-se eficaz para identificar padrões de retenção, reprovação acumulada e baixa aderência ao fluxo curricular entre discentes evadidos. Os resultados evidenciam o potencial da mineração de processos como ferramenta de apoio à gestão acadêmica, permitindo identificar disciplinas críticas, trajetórias de risco e oportunidades de intervenção institucional voltadas à permanência estudantil.

6. Conclusão

Este trabalho apresentou uma modelagem da estrutura curricular do curso de Ciência da Computação da Universidade Federal do Ceará (UFC), campus Quixadá, utilizando redes de Petri para representar formalmente trajetórias acadêmicas e aplicar técnicas de mineração de processos na análise da evasão estudantil. A partir dos históricos acadêmicos dos discentes, foi estruturado um *log* de eventos que possibilitou analisar a aderência entre os percursos realizados pelos estudantes e o fluxo curricular previsto.

Os resultados evidenciaram o potencial da técnica de *token replay* na identificação de padrões de retenção, reprovações acumuladas e trajetórias associadas à evasão. Além disso, a abordagem permitiu identificar disciplinas críticas e ramificações curriculares com maior concentração de gargalos acadêmicos, fornecendo subsídios para ações institucionais voltadas ao acompanhamento estudantil e à permanência no ensino superior.

Como limitação, o estudo considerou exclusivamente dados acadêmicos estruturados, sem incorporar fatores socioeconômicos, emocionais ou comportamentais relacionados à evasão. Além disso, a análise foi realizada em um único curso de graduação, embora a abordagem proposta possa ser adaptada para diferentes estruturas curriculares.

Como trabalhos futuros, pretende-se expandir a análise para outros cursos de graduação da UFC e explorar técnicas de *Object-Centric Process Mining* (OCPM), possibilitando uma análise integrada de múltiplos objetos acadêmicos envolvidos nas trajetórias estudantis.

7. Declaração sobre uso de Inteligência Artificial

Os autores declaram ter utilizado modelos de Inteligência Artificial Generativa (GPT-4.1), no apoio à revisão linguística deste trabalho. As ferramentas foram aplicadas com foco na melhoria da clareza, organização e coesão textual. Todas as decisões de conteúdo e autoria são de responsabilidade exclusiva dos autores.

Referências

- Carvalho, C., Mattos, J., and Aguiar, M. (2023). Avaliação da interpretabilidade de modelos por meio da clusterização de explicações no contexto da predição de evasão no ensino superior. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1191–1201, Porto Alegre, RS, Brasil. SBC.
- Colpo, M., Primo, T., and Aguiar, M. (2021). Predição da evasão estudantil: uma análise comparativa de diferentes representações de treino na aprendizagem de modelos genéricos. In *Anais do XXXII Simpósio Brasileiro de Informática na Educação*, pages 873–884, Porto Alegre, RS, Brasil. SBC.
- de Freitas Filho, A. G. (2025). Análise da trajetória acadêmica estudantil utilizando mineração de processos com ênfase na evasão do curso de Ciência da Computação da UFC em Quixadá. Dissertação (mestrado em computação), Campus de Quixadá, Programa de Pós-Graduação em Computação, Universidade Federal do Ceará, Quixadá.
- do Carmo, Ê. P., Gasparini, I., and de Oliveira, E. H. T. (2022). Identificação de trajetórias de aprendizagem em um curso de graduação e sua relação com a evasão escolar. In

- Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 323–333. SBC.
- dos Santos Oliveira, R. and de Medeiros, F. P. A. (2024). Modelo de predição de evasão escolar com base em dados de autoavaliação de cursos de graduação. *Revista Brasileira de Informática na Educação*, 32:1–21.
- INEP, D. d. E. E. (2017). Metodologia de cálculo dos indicadores de fluxo da educação superior.
- Jesus, P. A. S. d., Rodrigues, M. d. S. T., and Torres, J. B. (2014). Proposta de mapeamento de processos de armazenagem em centro de distribuição de supermercados de fortaleza/ce utilizando bpmn. Disponível em: <http://www.repositorio.ufc.br/handle/riufc/61995>. Acesso em: 11/06/2024.
- Kantorski, G. Z., Martins, R. Z., Balejo, A., and Frick, M. (2023). Mineração de dados educacionais para predição da evasão em cursos de graduação presenciais no ensino superior. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1133–1142. SBC.
- Morais, F., Melo, A., Moutinho, M., and Fagundes, R. (2021). Modelos de regressão aplicados na previsão da evasão escolar do ensino básico: uma revisão sistemática da literatura. In *Anais do XXXII Simpósio Brasileiro de Informática na Educação*, pages 168–178, Porto Alegre, RS, Brasil. SBC.
- Saraiva, D. V., Pereira, S. S., Braga, R. B., and de Oliveira, C. T. (2021). Análise de agrupamentos para caracterização de indicadores de evasão. In *Workshop sobre Educação em Computação (WEI)*, pages 238–247. SBC.
- Souza, O., Moraes, P., and Júnior, F. S. (2015). Um estudo sobre a evasão no curso de licenciatura em informática do ifrn–campus natal–zona norte. In *Anais do XXIII Workshop sobre Educação em Computação*, pages 216–225. SBC.
- Unger, A. J., Junior, D. A., Lima, F. O., Geraldo, I. C., Venero, S. K., and Ambrosio, R. R. (2022). Análise da trajetória de aprendizagem do aluno no ensino a distância por meio da mineração de processos. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 1163–1172. SBC.
- Van der Aalst, W. (2016). *Data Science in Action*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Viana, F., Santana, A., and Rabêlo, R. (2022). Avaliação de classificadores para predição de evasão no ensino superior utilizando janela semestral. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 908–919, Porto Alegre, RS, Brasil. SBC.
- Êrica Carmo, Gasparini, I., and Oliveira, E. (2022). Identificação de trajetórias de aprendizagem em um curso de graduação e sua relação com a evasão escolar. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 323–333, Porto Alegre, RS, Brasil. SBC.