

## Análise dos Resultados de Insucesso Escolar com o suporte de Mineração de Processos Educacionais

Edson P. Pimentel<sup>1</sup>, Eduardo M. Real<sup>1,2</sup>, Juliana C. Braga<sup>1</sup>, Wagner T. Botelho<sup>1</sup>

<sup>1</sup>Universidade Federal do ABC (UFABC)  
Santo André – SP – Brasil

<sup>2</sup>Universidade Estadual de Mato Grosso do Sul (UEMS)  
Nova Andradina – MS – Brasil

{edson.pimentel,eduardo.real,juliana.braga,wagner.tanaka}@ufabc.edu.br

**Abstract.** *The analysis of the students' learning results, based on the assessment instruments, is essential to feedback the teaching-learning process. However, data from summative assessments alone are not sufficient to understand student failures. The application of Educational Process Mining methods given in course event logs can reveal the learning trajectory and support a more robust analysis. This article aims to analyze the possible impacts of learning paths on students' performance failing a subject, conducted in a blended learning manner, that is, with classes and exercises at a distance and with face-to-face assessments.*

**Resumo.** *A análise dos resultados de aprendizagem dos estudantes, a partir dos instrumentos de avaliação, é essencial para retroalimentar o processo de ensino-aprendizagem. No entanto, somente os dados de avaliações somativas não são suficientes para compreender os insucessos dos estudantes. A aplicação de métodos de Mineração de Processos Educacionais em dados de log de eventos dos cursos podem revelar a trajetória de aprendizagem e apoiar uma análise mais robusta. Este artigo tem por objetivo analisar os possíveis impactos dos trajetos de aprendizagem no desempenho dos estudantes reprovados em uma disciplina, conduzida de forma semipresencial, ou seja, com aulas e exercícios a distância e com avaliações presenciais.*

### 1. Introdução

Este trabalho situa-se no contexto da educação mediada por Sistemas de Gerenciamento da Aprendizagem (*Learning Management Systems* - LMS), nas modalidades presencial, semipresencial ou a distância. Nesses sistemas, normalmente, o professor organiza a sequência do curso (conteúdos, atividades e comunicação) de acordo com suas estratégias didático-pedagógicas para guiar o aluno no seu processo de aprendizagem no curso. No entanto, a menos que restrições sejam impostas, os LMS possibilitam que os alunos possam optar por seguir caminhos diferentes daqueles indicados na organização do curso. É claro que isto pode ser inclusive parte da intencionalidade pedagógica do professor.

De acordo com Negri (2016), “a intencionalidade pedagógica é toda a ação consciente, planejada e executada pelo educador, acomodada dentro do cenário pedagógico (salas de aula ou qualquer outro ambiente no qual seja possível o ato de ensino e aprendizagem), determinado como espaço relacional dos que ensinam e dos que aprendem”.

De qualquer forma, o professor precisa de elementos para acompanhar os processos de aprendizagem dos estudantes. No entanto, somente os dados de avaliações somativas não são suficientes para compreender os sucessos e insucessos dos estudantes. Compreender os insucessos é essencialmente importante para retroalimentar o processo de ensino-aprendizagem e, se for o caso, reorganizar a sequência do curso, os materiais e os processos avaliativos.

Nos LMS a interação do estudante com os materiais do curso pode ser armazenada nos chamados registros de eventos (*logs*). Esses dados de *logs* (*logs* de eventos) têm o potencial de revelar de uma forma completa (início-fim) o comportamento do estudante a partir de como os conteúdos e atividades foram acessados e realizados por ele, i.e., os passos realizados durante os estudos nos ambientes LMS. Essa revelação pode ser alcançada com o uso de abordagens de técnicas de Mineração de Processos (MP), que têm como base de aplicação as evidências nos *logs* de eventos extraídos de sistemas de informação. No domínio da Educação, a aplicação da MP é conceituada como Mineração de Processos Educacionais (MPE). A ideia básica da MP é extrair conhecimento dos *logs* de eventos registrados por um sistema de informações. Assim, busca o confronto entre tais *logs* de eventos, i.e., comportamento observado e modelos de processos elaborado por especialista ou descobertos automaticamente (AALST, 2016).

Este artigo tem por objetivo analisar os possíveis impactos dos trajetos de aprendizagem no desempenho dos estudantes reprovados em uma disciplina semipresencial, conduzida por meio do LMS Moodle, com aulas e exercícios a distância e com avaliações presenciais. O artigo está organizado da seguinte forma: a seção 2 posiciona os conceitos fundamentais de MP e MPE e descreve alguns trabalhos relacionados; a seção 3 descreve os métodos utilizados para a condução dos experimentos; a seção 4 apresenta os resultados e discussões obtidos; e a seção 5 estabelece as considerações finais.

## 2. Revisão de Literatura

Esta seção tem por objetivo abordar os principais conceitos e formalismos necessários para a compreensão deste trabalho, bem como alguns trabalhos relacionados.

### 2.1. Mineração de Processos

A MP visa explorar os *logs* de eventos de maneira significativa para fornecer informações, identificar gargalos, antecipar problemas, recomendar contramedidas, otimizar processos etc. Para isso, foca em dados que permitem trilhar tempo e causalidade de atividades, obtendo eventos que são imperceptíveis a especialistas e não são tratados pelos algoritmos de AM comuns (AALST, 2016).

Formalmente, a entrada para a MP são os *logs* de eventos gerados nos sistemas de informação. Cada registro do *log* está relacionado a um evento e cada evento corresponde a uma atividade realizada, sendo que o *log* de eventos deve apresentar uma relação sequencial entre os eventos. Além disso, podem existir outras informações adicionais nos *logs* que também podem ser úteis, como data e hora (início e/ou fim), recurso (pessoa ou dispositivo que realizou), valores diversos (moedas, pesos, idades, status etc), entre outros. Os *logs* de eventos são descritos e definidos por alguns atributos básicos, sendo os dois primeiros o mínimo necessário para se obter um modelo, a saber:

- *Case* – elementos definidos como objetos centrais de análises. São organizados em ordem na sequência do *log* de eventos;

- *Activity Name* – atividades realizadas e que geram eventos no *log* de eventos;
- *Timestamp* – data e hora da gravação de eventos de *log* (início e/ou fim);
- *Resource* – identifica elementos que foram utilizados para executar as atividades (por exemplo, uma pessoa ou dispositivo).

Os passos de atividades realizados em cada *Case* forma um *Trace*, o qual é descrito por um sequenciamento ordenado de eventos (atividades), considerando que cada evento é único e está relacionado a apenas uma atividade (*Activity Name*). Neste contexto, um *log* de eventos é descrito por  $N$  *Cases*  $C$  ( $log = C_1, C_2, C_3, \dots, C_N$ ), em que cada um pode gerar um *Trace*  $T$  com até  $M$  tipos de eventos  $E$  ( $T = EC_1^*, EC_2^*, EC_3^*, \dots, EC_M^*$ ), obviamente respeitando uma sequência de eventos realizados e de acordo com o *Control-Flow* gerado. Além disso, também podem ocorrer repetições (\*) de eventos, se houver loops ou caminhos de retornos. Por exemplo, veja o suposto trecho de *log* de eventos a seguir, em que há  $N$  *Cases* e seus respectivos *Traces* descritos por até oito tipos de eventos  $E$  (a, b, c, d, e, f, g e h):

- *Cases*: 1, 2, 3, 4, 5, ...,  $N$ ;
- *Traces*: [ $\langle a, b, c, d, e, h \rangle$ ,  $\langle a, d, c, e, g \rangle$ ,  $\langle a, c, d, e, f, b, d, e, g \rangle$ ,  $\langle a, d, b, e, h \rangle$ ,  $\langle a, c, d, e, f, d, c, e, f, c, d, e, h \rangle$ , ...,  $\langle a, b, c, d, e, f, g, h \rangle$ ].

Note que o primeiro *Trace*, que é  $\langle a, b, c, d, e, h \rangle$ , corresponde a uma sequência de atividades realizada pelo *Case* 1, caracterizada por  $a \gg b \gg c \gg d \gg e \gg h$ . A mesma lógica pode ser utilizada para representar a sequência dos demais *Traces*. O fato é que em um *log* de eventos real a quantidade de *Cases* pode ser grande e o tamanho dos *Traces* pode variar muito, podendo conter de 1 a centenas de atividades.

Para a aplicação da MP, existem três tipos principais de conjuntos de técnicas que podem ser usados a partir de logs de eventos e modelos (AALST, 2016):

- *Discovery* - técnica de descoberta que tem como entrada um *log* de eventos (um arquivo de *logs*) e produz um modelo de processo;
- *Conformance* - um modelo de processo existente é comparado com o *log* de eventos pertencente ao mesmo processo, possibilitando verificar se a realidade, conforme registrada no *log*, está em conformidade com o modelo e vice-versa.
- *Enhancement* - possibilita estender ou melhorar um modelo de processo existente usando informações sobre o processo real registrado em algum *log* de eventos.

Além dos três tipos de MP existentes, há diferentes perspectivas de aplicação que podem ser consideradas e adotadas (AALST, 2016):

- *Control-flow* - concentra-se na ordem das atividades e o objetivo da mineração é encontrar uma boa caracterização de todos os caminhos possíveis;
- *Organizational* - concentra-se nas informações sobre os recursos ocultos no *log*, o objetivo é estruturar a organização classificando as pessoas em termos de funções e unidades organizacionais;
- *Case* - concentra-se nas propriedades dos *Cases*;
- *Time* - preocupa-se com o tempo e a frequência dos eventos.

Para avaliar um modelo obtido por algoritmos de *Discovery*, algum critério de qualidade pode ser utilizado, a partir de medidas correspondentes a quatro dimensões, ou forças, cujos valores ficam entre 0 e 1. São elas: *Fitness*, *Precision*, *Generalization* e *Simplicity* (AALST, 2016).

## 2.2. Mineração de Processos Educacionais

A MPE possibilita mapear o comportamento dos estudantes com base nos seus trajetos ao acessar conteúdos e realizar atividades num curso por meio de um LMS. Assim a MPE pode ser usada por educadores para, por exemplo, entender melhor os hábitos de aprendizagem dos estudantes, os fatores que influenciam seu desempenho e as habilidades adquiridas, criando e analisando modelos de processos educacionais que representam o comportamento observado (DOLAK, 2019; SALAZAR-FERNANDEZ; SEPULVEDA; MUNOZ-GAMA, 2019; GHAZAL; IBRAHIM; SALAMA, 2017; BOGARÍN; CEREZO; ROMERO, 2018).

Os modelos descobertos pela MPE podem ser usados para entender melhor os processos educacionais subjacentes, detectar precocemente dificuldades de aprendizado, gerar recomendações aos estudantes, ajudar os estudantes com dificuldades de aprendizagem específicas, fornecer *feedback* aos estudantes, docentes ou pesquisadores etc (BOGARÍN; ROMERO; CEREZO, 2016). Além disso, a MPE permite examinar quais ações específicas os estudantes executaram e verificar a categoria das ações sobre determinadas atividades (CEREZO; BOGARÍN; ROMERO, 2019). É possível ainda realizar procedimentos de análise de conformidade, verificando se um comportamento previamente modelado corresponde ao comportamento observado (SILVA; SILVA; ARANHA, 2016).

A MPE, como método, pode abrir novas maneiras de analisar o comportamento de aprendizagem e de soluções de problemas dos estudantes (TÓTH; ROLKE HEIKO; GOLDHAMMER; BARKOW, 2017). No entanto, a implementação da MP pode ser considerada um desafio interdisciplinar, exigindo que especialistas no assunto, desenvolvedores de itens, psicometristas e cientistas da computação trabalhem juntos para extrair, agregar, modelar e interpretar os dados adequadamente. Quanto a isto, Romero et al. (2016) enfatizam corretamente que a compreensibilidade de um modelo é um objetivo central da educação devido à transferência de conhecimento que isso implica. A compreensão de gráficos, modelos e representações visuais para docentes e estudantes torna os resultados essenciais para monitorar o processo de aprendizagem e fornecer *feedback*.

## 2.3. Trabalhos Relacionados

Esta seção apresenta e discute alguns trabalhos relacionados a este estudo. Eles estão agrupados de acordo com suas finalidades: como explorar os de *logs* de eventos e como visualizar o comportamento dos estudantes e seus resultados.

Romero et al. (2016) organizaram um tutorial mostrando como explorar *logs* de eventos coletados no LMS Moodle e obter modelos de processos contendo caminhos e atividades mais ou menos frequentes em relação ao número de ocorrências de execução. Neste trabalho, os autores trouxeram como proposta para facilitar a interpretação de resultados pelos docentes a segmentação de um *log* de eventos, dado que grandes conjuntos de dados resultam em modelos complexos para um docente ou estudante interpretar.

Beemt, Buijs e Aalst (2018) também usaram a segmentação em grupos (semanas do curso) de um *log* de eventos de um curso do tipo MOOC (*Massive Open Online Courses*) a fim de explorar a relação entre o comportamento de aprendizagem e seu progresso. Os resultados da análise de grupos indicaram que estudantes que assistiram regularmente a vídeos sucessivos em lotes tiveram um melhor desempenho nos resultados de apren-

dizagem. No entanto, os resultados indicam que esse comportamento está muito mais relacionado à ordem de exibição de vídeos do que ao tempo real.

O trabalho realizado por Cerezo, Bogarín e Romero (2019), tratou de uma abordagem sobre a avaliação da aprendizagem autorregulada em grupos de estudantes, sobre *logs* do Moodle. Como resultado da aplicação da MPE, este trabalho obteve que, no geral, embora os estudantes aprovados não sigam exatamente as orientações dos docentes (requisitos), eles seguem a lógica de um processo de aprendizado autorregulado suficiente, em oposição aos que reprovaram. O trabalho descrito em Etinger (2020) buscou descobrir e analisar os padrões de comportamentos realizados por estudantes com notas mais altas em contraste com os de notas mais baixas.

Também aplicando MPE sobre *logs* de eventos do Moodle, os autores em Dolak (2019) mostraram que uma análise detalhada sobre o comportamento dos estudantes é possível por meio de estatísticas e casos individuais de sequências de eventos de estudantes. Neste sentido, padrões de variações de comportamentos de estudantes foram obtidos identificando, por exemplo, comportamento típicos de estudantes ativos ou inativos.

A abordagem trazida em Salazar-Fernandez, Sepulveda e Munoz-Gama (2019) foi a de que a análise de trajetórias educacionais utilizando técnicas de MP pode ajudar a explicar a relação entre uma sequência de resultados acadêmicos e o abandono tardio. Estes autores criaram modelos de trajetetos em cursos com altas taxas de reprovação, tendo como objetivo compreender o processo que leva ao abandono tardio de um determinado curso, por meio da análise de trajetetos dos estudantes, identificando semelhanças e diferenças.

Este trabalho assemelha-se a Romero et al. (2016) e Etinger (2020) que aplicaram MPE sobre conjunto de grupos (aprovados e reprovados). Quanto ao trabalho de Salazar-Fernandez, Sepulveda e Munoz-Gama (2019) a semelhança está em olhar também para os abandonos. As principais diferenças estão relacionadas a: (i) observar esses estudantes em dois momentos: antes e depois da primeira avaliação presencial e (ii) utilizar outras informações além da análise visual do modelo de processo.

### 3. Materiais e Métodos

Os métodos utilizados neste estudo incluem: (a) a aplicação de MPE sobre dados de um *log* de eventos coletados a partir de uma disciplina de “Introdução a Programação” em um curso de graduação de uma universidade pública e (b) a análise das trajetórias de aprendizagem de estudantes reprovados. Para apoiar a análise foram adicionados aos dados do *log* notas obtidas pelos estudantes em 2 avaliações presenciais e também a classificação final de cada estudante: aprovado, reprovado, desistente.

A disciplina foi conduzida no LMS Moodle durante 12 semanas, na modalidade semipresencial com a participação inicial de 83 estudantes. Os estudantes tiveram o acompanhamento de 3 professores e 3 tutores, realizaram 43 atividades com entregas semanais obtendo *feedback* de todas as entregas. Além de assistir videoaulas e estudar os textos cada estudante deveria realizar atividades do tipo *quiz* (múltipla escolha) e também atividades de programação com correção automática ou manual (Laboratório de Programação Virtual). Os procedimentos de avaliação incluíram a realização de 2 provas presenciais, denominadas de P1, P2 e REC (recuperação).

O arquivo de *log* obtido do Moodle, relacionados a essa turma, possuía original-

mente 85759 registros de eventos descritos por sete atributos, a saber:

- IP Address: endereço IP (atributo desconsiderado).
- Timestamp: data e hora inicial do evento.
- Student: nome do estudante, ou seja, identifica o usuário que executou a ação.
- Component: especifica o tipo-atividade.
- Event Context: contém o título do tipo-atividade acessado.
- Event Name: indica a ação executada no tipo-atividade.

Os atributos Component, Event Context e Event Name permitem visualizar uma hierarquia, conforme expressa na Tabela 1, que apresenta dois exemplos. O nível 1 corresponde ao atributo Component, ou seja, as categorias de atividades organizadas para a disciplina: Laboratório de Programação Virtual (LPV), Lição, Questionário, etc. O nível 2 indica o Event Context, ou seja, o item específico acessado pelo estudante, por exemplo, "S01-04A- Estruturas Sequenciais(ex1)". Por fim, o nível 3 está relacionado ao Event Name, ou seja, com a ação específica do estudante no item: editar, enviar, etc.

**Tabela 1. Exemplos dos níveis de hierarquia.**

Nível	Atributo	Exemplo 1	Exemplo 2
1	Component	VPL	Lição
2	Event Context	S01-04A-Sequenciais	S03-01-Repetição(While-Do)
3	Event Name	edited, evaluated	viewed

De fato, essa hierarquia está relacionada às sequências de eventos ao projetar o curso. Primeiro, o professor seleciona o tipo de atividade (Component - nível 1) e depois define a atividade com seu conteúdo específico dando-lhe um nome (Event Context - nível 2). Finalmente, quando o estudante executa as ações na atividade, o evento é de fato registrado (Event Name - nível 3). A compreensão exata dos três níveis é importante, pois alguns ou muitos desses eventos de nível 3 podem não ser críticos para a análise e podem ser tratados na etapa de preparação e limpeza de dados.

Na etapa de limpeza de dados o atributo IP Address foi removido, assim como diversas ocorrências que não se mostraram necessárias no estudo, tais como: (a) eventos de nível 1 que não estavam relacionados a atividades de estudo ou avaliação, como por exemplo formulários de solicitação de prova substitutiva e Bate-Papo; (b) eventos de nível 3 relacionados a detalhes de ações que não estavam relacionados aos estudos ou avaliação; (c) eventos de nível 1 ou 3 relacionados a registros automáticos do Moodle.

A ordem das remoções iniciou pelo atributo Component, seguido por Event Context e Event Name. Assim, o *log* de eventos original sofreu uma redução resultando em 24334 eventos de 73 estudantes. Para os fins desse estudo extraiu-se um subconjunto com os dados dos 29 estudantes reprovados, contendo 3966 eventos para processamento e análise. Além disso, para apoiar outras análises, 4 segmentos foram extraídos do subconjunto de reprovados, usando-se os conceitos obtidos na primeira avaliação (P1): (i) 562 eventos dos 5 estudantes que obtiveram conceitos A ou B na P1; (ii) 453 eventos dos 5 estudantes que obtiveram conceitos C na P1; (iii) 756 eventos dos 6 estudantes que obtiveram conceitos D na P1; e (iv) 939 eventos de 13 estudantes que não realizaram a prova P1.

Escolheu-se a prova P1 como um elemento separador dos subconjuntos a serem analisados por ser a primeira avaliação presencial, e por abordar conteúdos mais elementares da disciplina como Sequência, Condição e Repetição. Os conteúdos após a prova P1 são Vetores, Matrizes e Modularização. Quanto à segmentação por conceitos obtidos na P1 temos: “A/B” equivale a aproveitamento superior a 80%; “C” equivale a aproveitamento superior a 65%; “D” equivale a aproveitamento superior a 45%; “NC” representa os que não compareceram à prova P1.

Para este estudo, a configuração do *log* de eventos foi definida pela associação do atributo Student como “Case” e do atributo Component como “Activity Name”. Além disso, as perspectivas de aplicação da MPE ocorreram sobre *Case*, *Control-Flow* (representados pelos modelos de processo) e *Time* (especificamente quanto à frequência das atividades). Para a mineração de processos utilizou-se o algoritmo do tipo *Discovery* denominado *Heuristic Miner* (Weijters, Aalst e Medeiros (2006) e Aalst (2016)), a partir de *scripts* desenvolvidos em Python com base na biblioteca PM4Py. Além disso, foram aplicadas técnicas para obter as informações estatísticas do *log* de eventos (tanto via *ProM Framework* quanto *scripts* Python desenvolvidos).

#### 4. Experimentos e Análises de Resultados

Apresentamos a seguir a descrição dos experimentos e respectivas análises sobre os 5 subconjuntos. De modo geral, os eventos presentes no *log* do curso estão relacionados majoritariamente a 3 tipo-atividades (atributo *Activity Name*), a saber: Laboratório de Programação Virtual (LPV), Questionário e Lição. É importante destacar que LPV e Questionário eram obrigatórias (avaliação), pois havia um percentual mínimo a ser enviado. Mesmo não sendo obrigatória, a intencionalidade pedagógica do professor almejava que o estudantes iniciassem o trajeto pelo tipo-atividade Lição. Ou seja, em tese, primeiro o estudante deveria estudar a lição e depois realizar as atividades de avaliação.

O modelo de processos do subconjunto completo dos 29 estudantes reprovados está representado na figura 1. O nó origem (verde) indica que dos 29 estudantes, 45% (13) iniciaram suas ações pelo tipo-atividade “Lição” e 55% (16) iniciaram por uma atividade avaliativa: LPV (4) e Questionário (12). Nela é possível visualizar também as conexões entre os nós tipo-atividades por meio de setas direcionadas entre eles, além dos valores que expressam a quantidade de ocorrências. Por exemplo, partindo do nó “Lição” 79 eventos seguiram para LPV, 59 para Questionário e 222 prosseguiram para outra Lição.

A Tabela 2 apresenta um resumo da distribuição de ocorrências de eventos dos 3 tipo-atividades mais representativos, em termos de quantidade de eventos. A primeira coluna (#N1) representa a quantidade de eventos dos 29 estudantes reprovados (*Cases*), por tipo-atividade. Nesta primeira coluna, as atividades avaliativas (LPV+Questionário) representam 90% ( $1805+1782=3587$ ) dos 3966 eventos realizados pelos 29 estudantes em 43 atividades possíveis. A primeira linha (*Cases*) representa o número de estudantes em cada subconjunto). A última linha (#EVENTOS) representa a quantidade de eventos em cada subconjunto.

Os demais valores, nas colunas dos subconjuntos, representam a média de eventos por estudante, em cada tipo-atividade. Os valores em negrito nas colunas #S1 e #S2 representam as médias totais, respectivamente, antes e após a prova P1. Os elementos sublinhados, indicam os valores acima da média total. Por exemplo nota-se na quarta



A primeira linha apresenta o número de estudantes que iniciaram seus trajetos por “Lição” ou por “Avaliação”: O período S01-S04 foi iniciado por 44 aprovados e 29 reprovados. Já o período S05-S09 foi iniciado por 42 de 44 aprovados e por 15 de 29 reprovados. Ressalta-se que 2 aprovados ( $44-42=2$ ) iniciaram por “Fórum” que não consta na tabela. Dos 16 estudantes que realizaram a P1, 1 deles já nem iniciou a segunda parte da disciplina.

**Tabela 3. Origem dos Trajetos de Aprendizagem por período da disciplina**

Início por	Aprovados		Reprovados					
	Período		Período		Resultados da P1			
	S01-S04	S05-S09	S01-S04	S05-S09	A/B	C	D	Total
Estudantes	44	42	29	15	5	5	6	16
Lição	<u>57%</u>	<u>52%</u>	45%	33%	<u>60%</u>	40%	50%	8
Avaliação	43%	48%	<u>55%</u>	<u>67%</u>	40%	<u>60%</u>	50%	8

A Tabela 3 destaca (vide os valores sublinhados) os maiores valores por onde cada subgrupo iniciaram seus trajeto. Nota-se que os aprovados iniciaram seus trajetos majoritariamente por “Lição”, o que de certa forma era a intenção do professor (57% no período S01-S04 e 52% em S05-09). Quanto aos reprovados, estes iniciaram seus trajetos majoritariamente por “Avaliação”: 55% no período S01-S04 e 67% em S05-09. Isto de certa forma mostra uma diferença em relação aos dois grupos. Uma possível explicação para que os reprovados tenham iniciado seus trajetos por Avaliação está associado aos prazos de entrega. Ou seja, estes estudantes podem não ter feito um bom planejamento de estudos e tiveram que ir diretamente para Avaliação para cumprir prazos de entrega.

Em relação aos “Resultados da P1” (subconjunto de reprovados) nota-se que os que obtiveram os melhores conceitos na P1 (A ou B), 60% iniciaram o trajeto por “Lição”, conforme intenção do professor. Já os que obtiveram conceito C (intermediário) 60% iniciaram por Avaliação.

## 5. Considerações Finais

Este estudo teve por objetivo analisar possíveis relações entre trajetos de aprendizagem e desempenho dos estudantes reprovados em uma disciplina semipresencial conduzida no LMS Moodle, aplicando-se Mineração de Processos Educacionais no *log* de eventos desses estudantes incluindo dados de avaliações presenciais (prova P1) e o resultado final.

Os principais achados mostram que: (a) os que obtiveram conceitos altos na P1 tiveram média de eventos maior em LPV indicando que praticaram mais; (b) os que obtiveram o menor conceito na P1 tiveram média de eventos maior em Questionário e nem tanto em LPV, antes da P1; (c) os aprovados, em sua maioria, iniciaram seus trajetos por “Lição”, o que era esperado pelo professor, diferentemente dos reprovados que, em sua maioria, iniciaram seus trajetos por “Avaliação”, principalmente na segunda parte do curso. Assim, as análises apontam diferenças entre os grupos, mesmo que não sejam tão expressivas ou conclusivas.

Como trabalhos futuros essa pesquisa pretende projetar uma arquitetura para facilitar o trabalho do professor na análise desses trajetos de aprendizagem, especialmente dos estudantes que apresentam desempenhos ruins no decorrer do curso.

## Referências

- AALST, W. van der. *Process Mining: Data Science in Action*. 2. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 2016.
- BEEMT, A. van den; BUIJS, J.; AALST, W. van der. Analysing structured learning behaviour in massive open online courses (moocs): An approach based on process mining and clustering. *International Review of Research in Open and Distance Learning*, Athabasca University, v. 19, n. 5, p. 38–60, 2018.
- BOGARÍN, A.; CERESO, R.; ROMERO, C. Discovering learning processes using inductive miner: A case study with learning management systems (lms). *Psicothema*, p. 322–329, 08 2018.
- BOGARÍN, A.; ROMERO, C.; CERESO, R. Applying data mining to discover common learning routes in moodle. *Revista Edmetic*, p. 73–92, 2016.
- CERESO, R.; BOGARÍN, A.; ROMERO, C. Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education, Springer Science+Business Media, LLC, part of Springer Nature*, p. 1–15, 2019.
- DOLAK, R. Using process mining techniques to discover student’s activities, navigation paths, and behavior in lms moodle. In: RØNNINGSBAKK, L. et al. (Ed.). *Innovative Technologies and Learning*. [S.l.]: Springer International Publishing, 2019. p. 129–138.
- ETINGER, D. Discovering and mapping lms course usage patterns to learning outcomes. In: SPRINGER. *International Conference on Intelligent Human Systems Integration*. [S.l.], 2020. p. 486–491.
- GHAZAL, M. A.; IBRAHIM, O.; SALAMA, M. A. Educational process mining: A systematic literature review. In: *2017 European Conference on Electrical Engineering and Computer Science (EECS)*. [S.l.: s.n.], 2017. p. 198–203.
- NEGRI, P. S. A intencionalidade pedagógica como estratégia de ensino mediada pelo uso das tecnologias em sala de aula. *Laboratório de Tecnologia Educacional*, 2016.
- ROMERO, C. et al. Educational process mining: A tutorial and case study using moodle data sets. *Data Mining And Learning Analytics: Applications in Educational Research*, John Wiley Sons, n. 1, p. 3–27, 2016.
- SALAZAR-FERNANDEZ, J.; SEPULVEDA, M.; MUNOZ-GAMA, J. Influence of student diversity on educational trajectories in engineering high-failure rate courses that lead to late dropout. In: . [S.l.: s.n.], 2019. p. 607–616.
- SILVA, F.; SILVA, T. R.; ARANHA, E. Mineração de processo educacional - uma revisão sistemática da literatura. *Brazilian Symposium on Computers in Education*, v. 30, p. 1036–1045, 2016.
- TÓTH, K.; ROLKE HEIKO; GOLDHAMMER, F.; BARKOW, I. Educational process mining. new possibilities for understanding students’ problem-solving skills. In: \_\_\_\_\_. [S.l.]: The Nature of Problem Solving: Using Research to Inspire 21st Century Learning - OCDE Publishing, 2017.
- WEIJTERS, A.; AALST, W.; MEDEIROS, A. *Process Mining with the Heuristics Miner-algorithm*. [S.l.: s.n.], 2006. v. 166. 1-34 p.