

O uso de Learning Analytics no Brasil: uma revisão sistemática da literatura

Pedrina Célia Brasil¹⁹

Tainá Jesus Medeiros²⁰

Isabel Dillmann Nunes²¹

RESUMO

Evidências mostram que o uso da análise de aprendizagem ainda está em seus estágios iniciais. No Brasil os trabalhos neste campo de pesquisa ainda são muito recentes e escassos. O propósito deste trabalho é apresentar os resultados de uma Revisão Sistemática da Literatura (RSL) sobre estudos que abordam/relatam o uso de análise de aprendizagem em organizações brasileiras. Ao total 321 artigos foram identificados, destes 17 foram considerados relevantes ao propósito desta pesquisa.

Palavras-chave: Learning Analytics. Ambientes Virtuais de Aprendizagem. Revisão Sistemática da Literatura

1. Introdução

A adoção em larga escala de Ambientes Virtuais de Aprendizagem (AVA) — também conhecidos como Sistemas de Gerenciamento de Aprendizagem (SGA) — significa hoje que instituições de ensino e educadores podem tirar vantagem dos conjuntos de dados extraídos e armazenados através desses sistemas (ALMAZROUI, 2013).

Os SGAs normalmente possuem características tais como: compartilhamento de conteúdo pedagógico; acompanhamento de progresso acadêmico; opções de comunicação entre usuários (fóruns, chats, *etc*); elaboração e execução de atividades avaliativas; e entre outros (ROMERO et al., 2013). Embora esses sistemas promovam inúmeras funcionalidades, eles ainda carecem de ferramentas para melhor agrupamento, relato, visualização e análise de aprendizagem dos usuários (DAWSON, 2009). Professores reclamam da falta de funções que apoiem a tomada de decisão docente, indicando, por exemplo, a qualidade do aprendizado dos

¹⁹ Programa de Pós-Graduação em Inovação e Tecnologias Educacionais – PPGITE, Universidade Federal do Rio Grande do Norte – UFRN, 1524 – Campus Universitário Lagoa Nova – CEP 59072-970 – Natal-RN, Brasil. pedrina.brasil@gmail.br

²⁰ Escolas de Engenharias e Tecnologia de Informação - Universidade Potiguar – UnP 59082-902, Av. Engenheiro Roberto Freire, 2184 - Capim Macio, Natal - RN, Brasil. taina.medeiros@unp.br.

²¹ Programa de Pós-Graduação em Inovação e Tecnologias Educacionais – PPGITE, Universidade Federal do Rio Grande do Norte – UFRN, 1524 – Campus Universitário Lagoa Nova – CEP 59072-970 – Natal-RN, Brasil. bel@imd.ufrn.br

alunos; se as atividades são suficientemente desafiadoras; se os alunos estão motivados, chateados ou confusos; e *etc* (DRINGUS e ELLIS, 2005).

Nesse contexto, surge o campo de *Learning Analytics* (LA) — em português, Análise de Aprendizagem. LA é definida como a medida, coleta, análise e relato dos dados sobre os alunos e seus contextos de aprendizagem com o objetivo de entender e otimizar o aprendizado e o ambiente que ele ocorre (FERGUSON, 2012). Por meio da Análise da Aprendizagem é possível otimizar as oportunidades de um processo de ensino personalizado ao contexto do usuário, uma vez que a análise possibilita a descoberta das dificuldades de aprendizagem que o aluno enfrenta no seu dia a dia escolar.

A Análise de Aprendizagem tem sido utilizada, por exemplo, para avaliar cursos *on-line*, dando suporte ao desenvolvimento de sistemas de *e-learning* (ensino eletrônico) mais efetivos ao contexto dos alunos (BAKER e YACEF, 2009; KOTSIANTIS, 2012; SIEMENS e BAKER, 2012).

O propósito deste artigo é apresentar os resultados de uma Revisão Sistemática da Literatura (RSL) sobre estudos que abordam/relatam o uso de Análise de Aprendizagem. Este trabalho tem como objetivo identificar como a análise de aprendizagem tem sido aplicada a sistemas de apoio tecnológico para ambientes de ensino de instituições brasileiras.

Ao total foram identificados 321 trabalhos, relacionados automaticamente pelos motores de busca da SCOPUS²² e da Comissão Especial da Informática na Educação (CEIE)²³. Dos trabalhos identificados, 300 foram rejeitados a partir da aplicação dos critérios de inclusão e exclusão e avaliação de qualidade; 21 foram selecionados para análise de seleção inicial e final desta revisão. Dos 21 trabalhos selecionados, foi verificado que: 2 trabalhos são duplicados; e 2 trabalhos não possui texto na íntegra passível de acesso pela *web*. Desta forma, foram extraídos os dados de 17 trabalhos, que embasam as descobertas e discussões desta pesquisa.

Para contextualizar as decisões desta revisão, na Seção 2, está descrita a metodologia de pesquisa utilizada. Na Seção 3, é apresentado o protocolo de revisão adotado para a seleção de trabalhos e extração de seus dados. Na Seção 4, os resultados obtidos são apresentados e

²² <https://www.scopus.com/>

²³ www.br-ie.org/pub/index.php/index

confrontados com as afirmações encontradas na literatura. Por fim, são realizadas as considerações finais deste trabalho na Seção 5.

2. Metodologia do Trabalho

Este artigo apresenta os resultados de uma Revisão Sistemática da Literatura (RSL) que teve como objetivo analisar os trabalhos relacionados ao uso de Análise de Aprendizagem em ambientes educacionais brasileiros.

A Revisão Sistemática da Literatura (RSL), também conhecida como revisão sistemática, é um método de pesquisa que visa estabelecer um levantamento formal do estado da arte, de forma robusta e consistente, a partir de um planejamento e execução criteriosa de atividades (BIOLCHINI et al., 2005).

Segundo Kitchenham e Charters (2007), uma RSL emprega um processo metódico para identificar, avaliar e interpretar as evidências científicas disponíveis e relevantes relacionadas a um tema específico de pesquisa. A condução desse método é realizada a partir de um processo bem definido que envolve a execução de três etapas: Planejamento; Condução; e Publicação de Resultados (Tabela 1).

Tabela 1. Objetivos e etapas da RSL

Etapa	Objetivos	Resumo de Atividades
Planejamento	Definir o objetivo e planejar a Revisão Sistemática	<ul style="list-style-type: none"> ● Identificar a necessidade da Revisão Sistemática; ● Definir os objetivos da pesquisa; ● Criar o protocolo (planejar a RSL);
Condução (Identificação, seleção de estudos e extração de dados)	Executar o planejamento feito no protocolo, buscar estudos primários e selecionar os estudos a serem sintetizados	<ul style="list-style-type: none"> ● Executar as <i>strings</i> de busca nas máquinas de busca selecionadas; ● Selecionar os estudos primários de acordo com os critérios de inclusão e exclusão; ● Extrair informações dos estudos primários selecionados;
Publicação dos Resultados	Sintetizar os estudos primários que atendem ao propósito da revisão	<ul style="list-style-type: none"> ● Sintetizar as informações extraídas dos estudos primários; ● Publicar os resultados (relatório técnico ou artigos)

Adaptado de KITCHENHAM e CHARTERS, 2007.

O protocolo usado para realizar esta RSL foi baseado no trabalho de Kitchenham e Charters (2007). A principal diferença entre o método aqui utilizado e originalmente relatado é que, ao invés de utilizar um processo de busca manual, foi utilizada uma busca automatizada para identificação dos trabalhos candidatos a esta pesquisa.

Na etapa de planejamento foi identificada a necessidade de pesquisa e criado o protocolo de revisão deste trabalho. Segundo Kitchenham e Charters (2007), o protocolo é um documento que visa minimizar os vieses que possam ser cometidos pelos pesquisadores durante a execução de uma RSL. Neste trabalho, esse documento direcionou a condução do processo da revisão, tornando a mais confiável e passível de reprodutibilidade.

Na etapa de condução foram executadas as estratégias e processos planejados no protocolo de revisão. Nesta etapa, são utilizadas as bases bibliográficas SCOPUS e CEIE para identificação dos estudos relacionados ao objetivo desta pesquisa. A base SCOPUS foi selecionada por ser a maior base multidisciplinar de citações e resumos de literatura revisada por pares (periódicos, livros e conferências), segundo dados da ELSEVIER (2018). Uma estratégia utilizada neste trabalho foi combinar bases bibliográficas abrangentes, como a SCOPUS, com outras bases específicas. Dessa forma, a base CEIE (Comissão Especial de Informática na Educação) foi escolhida por reunir trabalhos de eventos nacionais relacionados especificamente a área de informática na educação, tais como: RBIE (Revista Brasileira de Informática na Educação), SBIE (Simpósio Brasileiro de Informática na Educação), WIE (Workshop de Informática na Escola), WCBIE (Workshops do Congresso Brasileiro de Informática na Educação), JAIE (Jornada de Atualização em Informática em Educação) e DesafiE.

Ainda na etapa de condução, os pesquisadores deram início ao processo de seleção dos estudos por meio da aplicação dos critérios de seleção (inclusão e exclusão) e avaliação de qualidade dos trabalhos. Os dados dos estudos selecionados foram extraídos e sintetizados de modo a facilitar as análises e sumarização dos resultados.

Na etapa de publicação, os resultados deste trabalho foram divulgados por meio de publicação de artigos.

3. Protocolo da Revisão Sistemática

Nesta seção são apresentados os critérios e estratégias utilizados na realização desta revisão sistemática. A partir desta, espera-se obter uma visão geral dos estudos propostos na literatura que reportam o uso de Análise de Aprendizagem em ambientes virtuais de ensino brasileiro, identificando e analisando as estratégias, métodos e resultados da aplicação dessa tecnologia.

3.1. Questões de Pesquisa

Esta revisão tem como questão central de pesquisa a seguinte pergunta: Como a análise de aprendizagem tem sido aplicada em ambientes virtuais de aprendizagem brasileiros? Para responder essa pergunta, foram definidas as seguintes questões secundárias:

- **QP1:** Que objetivos relacionados a análise de aprendizagem são abordados?
- **QP2:** Que instituições de ensino/pesquisa se utiliza de análise de aprendizagem?
- **QP3:** Quais dados são os mais explorados nas atividades de coleta?
- **QP4:** Que abordagens/tecnologias de coleta de dados são as mais aplicadas?
- **QP5:** Que abordagens/tecnologias de análise são as mais aplicadas?
- **QP6:** Os estudos apresentam resultados positivos as abordagens de análise de aprendizagem propostas?
- **QP7:** Que ambientes virtuais de aprendizagem se tem utilizado de análise de aprendizagem?

3.2. Processo de Pesquisa

A pesquisa foi realizada por meio de buscas automatizadas nos sistemas SCOPUS e CEIE. Todas as pesquisas foram baseadas: no título, palavras-chave e resumo dos trabalhos. As buscas ocorreram entre março e abril de 2018. Para todas as fontes, foi utilizado um conjunto de *strings* que agregou como resultante uma relação de trabalhos, conforme Tabela 2

Tabela 2. *Strings* de Busca Utilizadas

ID	Fonte	String	Resultado
S01	SCOPUS	("learning analytic" OR "learning analytics" OR "learning data mining" OR "educational data mining" OR "academic data mining") AND ("virtual learning environment" OR "learning management system")	311
S02	SCOPUS	("análise de aprendizagem" OR "mineração de dados educacionais" OR "mineração de dados acadêmicos" OR "mineração de dados escolares" OR "mineração de dados de aprendizagem") AND ("ambiente virtual de aprendizagem" OR "ambientes virtuais de aprendizagem")	0
S03	CEIE	("análise de aprendizagem" OR "mineração de dados educacionais" OR "mineração de dados acadêmicos" OR "mineração de dados escolares" OR "mineração de dados de aprendizagem") AND ("ambiente virtual de aprendizagem" OR "ambientes virtuais de aprendizagem")	8
S04	CEIE	("learning analytic" OR "learning analytics" OR "learning data mining" OR "educational data mining" OR "academic data mining") AND ("virtual learning environment" OR "learning management system")	2

Uma vez que a estratégia de busca adotada resultou trabalhos tanto da base SCOPUS quanto da base CEIE, foi utilizada a ferramenta Start²⁴, para melhor gerenciamento e controle dos estudos identificados. Todos os trabalhos foram registrados na ferramenta de forma manual e automática a partir da importação de arquivos de referência no formato BibTex²⁵.

3.3. Seleção dos Estudos

O processo de busca automatizada identificou 321 estudos (311 estudos da base SCOPUS; e 10 estudos na base CEIE). A seleção dos artigos aconteceu em 3 fases: 1. **Pré-Seleção**, 2. **Seleção Inicial** e 3. **Seleção Final**. Em cada uma dessas fases, os estudos identificados foram submetidos à análise dos critérios de inclusão e exclusão definidos nesta revisão conforme Tabela 3.

²⁴ http://lapes.dc.ufscar.br/tools/start_tool

²⁵ <http://www.bibtex.org/>

Tabela 3. Critérios de Inclusão e Exclusão de Estudos

Critérios de Inclusão	Critérios de Exclusão
<p>I1. Artigos publicados de janeiro/2012 a abril/2018; e</p> <p>I2. Artigos em português ou inglês; e</p> <p>I3. Artigos que propõem/relatam um processo, ferramenta, arquitetura ou diretriz para análise de dados de sistemas de apoio tecnológico ao ensino.</p>	<p>E1. Estudos que não atendem os critérios de inclusão; ou</p> <p>E2. Estudos não afiliados a instituições brasileiras de ensino e ou pesquisa; ou</p> <p>E3. Estudos duplicados; ou</p> <p>E4. Estudos sem resumo; ou</p> <p>E5. Estudos que não são primários; ou</p> <p>E6. Estudos que não foi possível acessar o texto na íntegra; ou</p> <p>E7. Estudos que não são artigos de revista/evento científico.</p>

Na fase de **pré-seleção**, os estudos identificados são relacionados e, a partir do tipo do estudo, afiliação dos autores, ano de publicação e idioma do texto, são analisados. Para não ser rejeitado nesta fase, os estudos precisam seguir os critérios de inclusão e exclusão definidos nesta pesquisa.

Em seguida, é iniciada a segunda etapa de seleção dos estudos, **seleção inicial**. Nesta etapa são aplicados os critérios de inclusão e exclusão em todos os estudos candidatos, por meio da avaliação do título, resumo e palavras-chave. Os estudos aqui aceitos seguem para a próxima etapa de seleção, conforme apresentado na Figura 1.

Na etapa de **seleção final** são aplicados os critérios de inclusão e exclusão dos estudos aceitos na seleção inicial, por meio da avaliação da introdução e conclusão dos trabalhos, conforme apresentado na Figura 1.

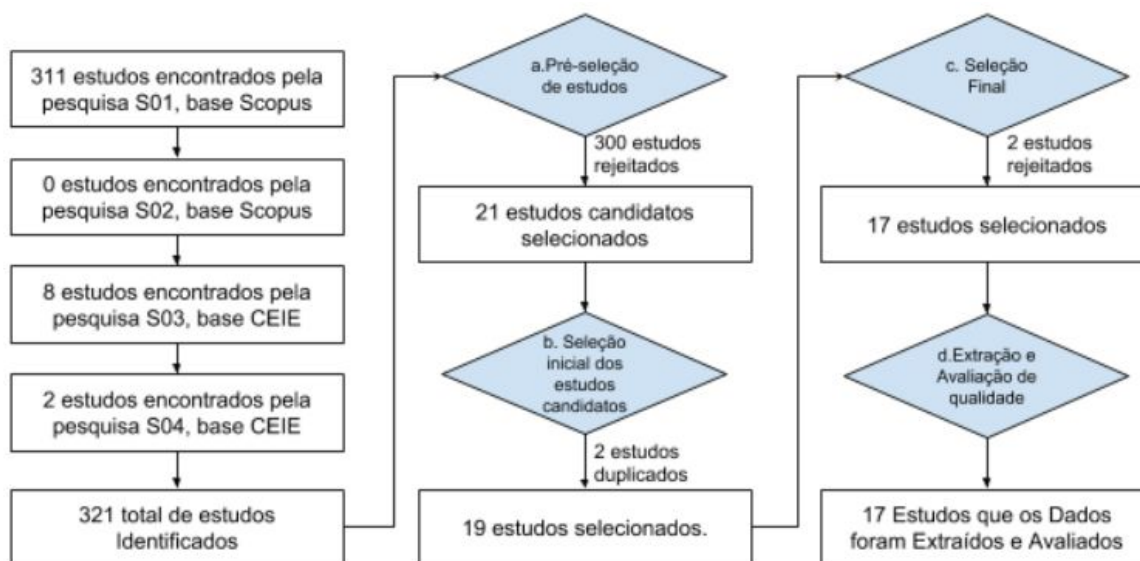


Figura 1. Processo de Seleção

Ao final de cada etapa de seleção, é realizada uma revisão dos estudos aceitos garantindo assim a confiabilidade e reprodutibilidade desta revisão.

3.4. Procedimentos de decisão

A partir da relação dos estudos identificados, os pesquisadores (R1 e R2) avaliam individualmente cada estudo, adicionando parecer de “aceito” ou “rejeitado” a cada um dos trabalhos avaliados. Cada pesquisador deverá avaliar todos os trabalhos de forma independente e individual, garantindo assim a confiabilidade do processo de seleção. No caso de dúvida, o pesquisador deverá, a princípio, marcar o estudo como aceito e como sendo obrigatória a sua revisão.

Em seguida, em um quadro de seleção geral, são relacionados os estudos selecionados e as respectivas avaliações de R1 e R2. As discordâncias devem ser solucionadas de forma colaborativa e consensual, através de revisões as avaliações realizadas. Após esse passo, as discordâncias restantes serão avaliadas por um terceiro avaliador externo R3, que as julgará produzindo um dos dois resultados: aceito ou rejeitado.

Em seguida, os estudos incluídos são integrados, por R1, numa lista final de estudos selecionados. Finalizados esses procedimentos, cada estudo selecionado deve ser lido completamente e a ele ser aplicado os critérios de avaliação de qualidade.

3.5. Avaliação de Qualidade

Os critérios de qualidade têm o objetivo avaliar os aspectos metodológicos dos estudos selecionados (NAKAGAWA, 2017). Através deles é possível, por exemplo, considerar os aspectos de relevância do tema do trabalho e o uso dos métodos que conduzam aos objetivos propostos.

Para esta revisão foram definidas 5 (cinco) questões para avaliação da qualidade dos trabalhos, conforme relacionadas a seguir:

- **QA1:** O estudo apresenta algum método, técnica ou ferramenta existente no mercado para auxiliar na sua abordagem?
- **QA2:** O estudo apresenta objetivos e/ou questões de pesquisa bem definidos e embasados pela literatura?
- **QA3:** O estudo avaliado apresenta algum tipo de experimento controlado ou estudo de caso para avaliação da abordagem proposta?
- **QA4:** O estudo avaliado apresenta uma breve comparação dos seus resultados com resultados de outros trabalhos relacionados?
- **QA5:** O estudo avaliado apresenta resultados extraídos de um contexto real de ensino?

Conforme o procedimento previsto em Kitchenham *et al.* (2010), para cada uma das questões, é possível uma das seguintes pontuações: 1 (Sim, o estudo apresenta); 0.5 (O estudo apresenta parcialmente); e 0 (Não, o estudo não apresenta). Ao final da avaliação, a pontuação dos estudos é calculada por adição aritmética das respostas. Os artigos com pontuação igual a 0 (zero) são rejeitados e eliminados da extração de dados.

3.6. Processo de Extração de Dados

Visando responder as questões desta pesquisa, foram definidos os seguintes atributos para extração de dados (Tabela 4).

Tabela 4. Extração de dados

Questões de Pesquisa	Atributo	Tipo de Dado
QP2: Que instituições de ensino/pesquisa se utiliza de análise de aprendizagem?	Título do trabalho.	Texto
	Afiliação dos autores.	Texto
QP1: Que objetivos relacionados a análise de aprendizagem são abordados?	Modalidade de curso (s) analisado (s).	Lista pré-definida (Presencial, Semipresencial, A Distância e Não Informado)
	Objetivo da análise de aprendizagem.	Texto
QP7: Que ambientes virtuais de aprendizagem se tem utilizado de análise de aprendizagem?	Ambiente (s) virtual (is) de ensino analisado (s).	Texto
QP3: Quais dados são os mais explorados nas atividades de coleta?	Dados coletados para análise de aprendizagem.	Texto
QP4: Que abordagens/tecnologias de coleta de dados são as mais aplicadas?	Abordagem e/ou ferramenta (s) de coleta de dados utilizado (s).	Texto
QP5: Que abordagens/tecnologias de análise são as mais aplicadas	Abordagem e/ou ferramenta (s) de análise de dados utilizada.	Texto
QP6: Os estudos apresentam resultados positivos as abordagens de análise de aprendizagem propostas?	Resultados relatados pelo estudo.	Lista pré-definida (Sim, Não, e Não Diagnosticado)

O processo de extração de dados e avaliação de qualidade dos estudos são realizados ao mesmo tempo, utilizando os procedimentos de avaliação de qualidade já descrito neste trabalho.

4. Resultados e Análise de Dados

Inicialmente foram identificados 321 artigos no total (311 SCOPUS e 10 CEIE). Seguindo os procedimentos de seleção, 300 estudos foram rejeitados na etapa de pré-seleção pois não atendiam totalmente os critérios de inclusão (Figura 2a) e atendiam total ou parcialmente os critérios de exclusão (Figura 2b).

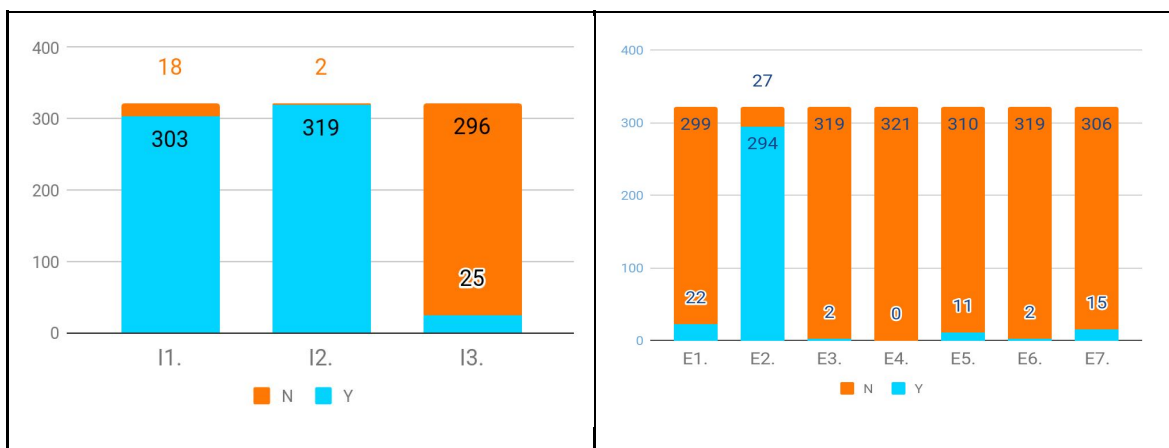


Figura 2a. Etapa de Pré-Seleção - Análise dos Critérios de Inclusão

Figura 2b. Etapa de Pré-Seleção - Análise dos Critérios de Exclusão

Os 21 estudos candidatos remanescentes da pré-seleção foram submetidos ao processo de seleção inicial e final. Individualmente cada pesquisador (R1 e R2) analisou os estudos candidatos aplicando os critérios de inclusão e exclusão a partir da leitura do título, resumo, palavras-chave, autores e afiliação dos trabalhos.

Na seleção inicial, 2 artigos foram considerados duplicados por ambos os pesquisadores, 2 artigos foram rejeitados por atenderem o critério de exclusão E6 (Estudos que não foi possível acessar o texto na íntegra). Os artigos remanescentes foram submetidos a etapa de seleção final. Nesta, os artigos foram analisados aplicando-se os critérios de inclusão e exclusão a partir da leitura do introdução, conclusão e afiliação dos trabalhos. Após a etapa de seleção final, o número total de artigos relevantes resultou em 17, os quais foram submetidos ao processo de extração de dados e avaliação de qualidade, conforme Tabela 5 disponível em: <https://goo.gl/fwKNZp>.

4.1. Respostas às Questões de Pesquisa e Análise de dados

Os 17 estudos (Tabela 5) foram incluídos no processo de extração de dados e avaliação de qualidade. Todos os estudos foram indexados com um código Identificador Único (Id) e, a partir da leitura completa desses trabalhos, foram realizadas as descobertas e análise das questões de pesquisa deste artigo.

QP1: Que objetivos relacionados à análise de aprendizagem são abordados?

Ao analisar os estudos, cada um deles aborda uma aplicação da análise de aprendizagem conforme objeto de estudo do trabalho. Somente 1 artigo (Id: 26956) não apresentou um objetivo fim a abordagem de análise de aprendizagem proposta. Cinco artigos (Ids: 26836, 26973, 27090, 26878 e 26961) propõem a aplicação da análise de aprendizagem na identificação de estudantes com baixa performance. Três artigos (Ids: 26836, 26973, 27094) propõem a aplicação da análise de aprendizagem na identificação de alunos em risco de evasão. Quatro artigos (Ids: 27048, 26877, 2114, 26961) propõem a aplicação da análise de aprendizagem para auxiliar a tomada de decisão docente através de de síntese aos dados de performance, comunicação e acessos do usuário. Outras aplicações de análise de aprendizagem identificadas foram: Acompanhar atividades de tutoria (Id: 26892); Avaliar a satisfação dos alunos em relação ao conteúdo, processo e resultados de aprendizagem (Id: 26878); Identificar grupos de alunos com características em comum (Ids: 26914, 27091); Identificar o comportamento das interações dos alunos em AVA (Ids: 27095, 2796), Promover participação efetiva dos alunos (Id: 26892); Identificar o comportamento dos alunos no uso da ferramenta de fórum (Id: 27092); Checar a proficiência gramatical dos alunos no uso da Língua Portuguesa-Br (Id: 27113); Assistir o estudante durante o processo de aprendizagem (Id: 26878); Projetar um modelo de avaliação autêntico, que possa fornecer avaliação e feedback aos alunos submetidos a metodologia PBL (Id: 26878); Entender os perfis de respostas para orientar as próximas atividades de aprendizagem (Id: 26961). Verificamos que 58% dos estudos tem como objetivo a aplicação da análise de aprendizagem em cursos na modalidade à distância.

QP2: Que instituições de ensino/pesquisa se utiliza de análise de aprendizagem?

Ao analisar que Instituições de Ensino/Pesquisa estão envolvidas com o uso de análise de aprendizagem no Brasil, foram encontradas 20 Instituições. Observamos que as Instituições em destaques se encontram no Nordeste, mais especificamente no Estado de Pernambuco. Verificamos que o Brasil em relação ao mundo possui pouquíssimos trabalhos relacionados ao uso de análise de aprendizagem em ambientes virtuais de ensino, representando somente 5% de todos os trabalhos identificados nesta pesquisa. Entre as instituições que apresentam trabalhos recorrentes nesta área de pesquisa em território nacional estão: UFRN (2), UFCG (2), UFPE (3), UPE (3) e UFRPE (3).

QP3: Quais dados são os mais explorados nas atividades de coleta?

Os dados mais explorados pelos estudos foram respectivamente: Número de postagem enviadas em fórum (utilizado em 47% dos estudos); Matrícula do Aluno (utilizado em 35% dos estudos), Número de Acessos ao Ambiente de Aprendizagem (utilizado em 29% dos estudos); Postagens enviadas em chats (utilizado em 29% dos estudos), Interações do Usuário no Ambiente (utilizado em 23% dos estudos), Identificação do Curso (utilizado em 23% dos estudos), Notas dos Alunos (utilizado em 23% dos estudos), Postagens de fóruns lidas pelos alunos (utilizado em 23% dos estudos).

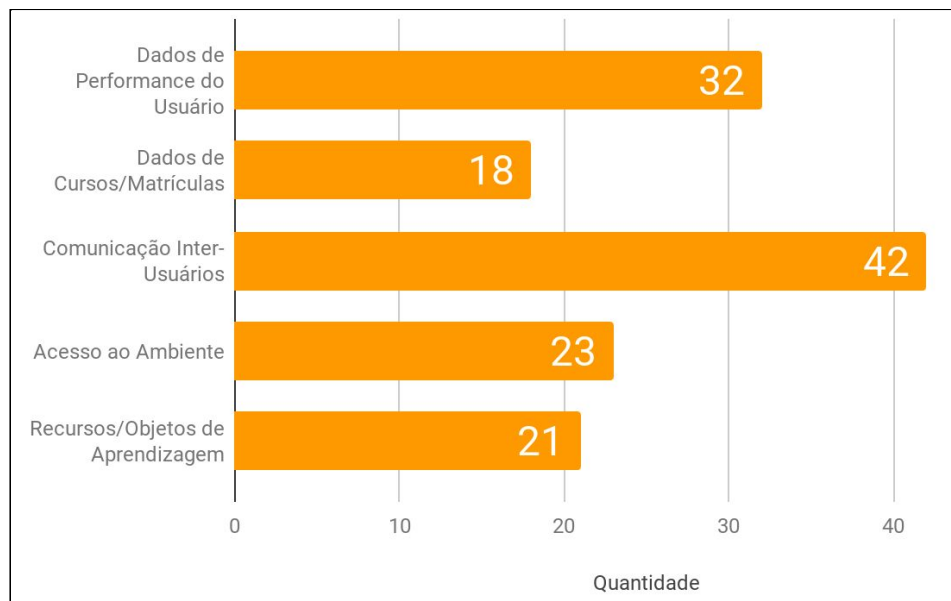


Figura 3. Gráfico referente aos dados mais coletados

Conforme ilustrado na Figura 3, percebe-se que a maioria dos dados utilizados nas análises de aprendizagem são dados quantitativos que dizem respeito a performance, identificação, vínculo com cursos e comunicação entre usuários.

QP4: Que abordagens/tecnologias de coleta de dados são as mais aplicadas?

A grande maioria dos trabalhos (70%) relatou as tecnologias utilizadas durante a coleta dos dados de suas pesquisas. Alguns trabalhos utilizam-se de mais de uma tecnologia de coleta (Ids: 26878, 26973, 27113, 26877 e 26956), outros somente uma tecnologia. As tecnologias de coleta mais utilizadas são: Consultas SQL à Base de Dados do SGA (utilizado por 23% dos estudos); Consultas à Arquivos de log do SGA (utilizado por 17% dos estudos) e Consultas SPARQL a arquivos de dados educacionais armazenados em formato OWL/RDF (utilizado 11%). Apesar dessa diversidade, verificamos que a utilização de arquivos de log e base de dados relacional encontram-se em mais uso (Figura 4).

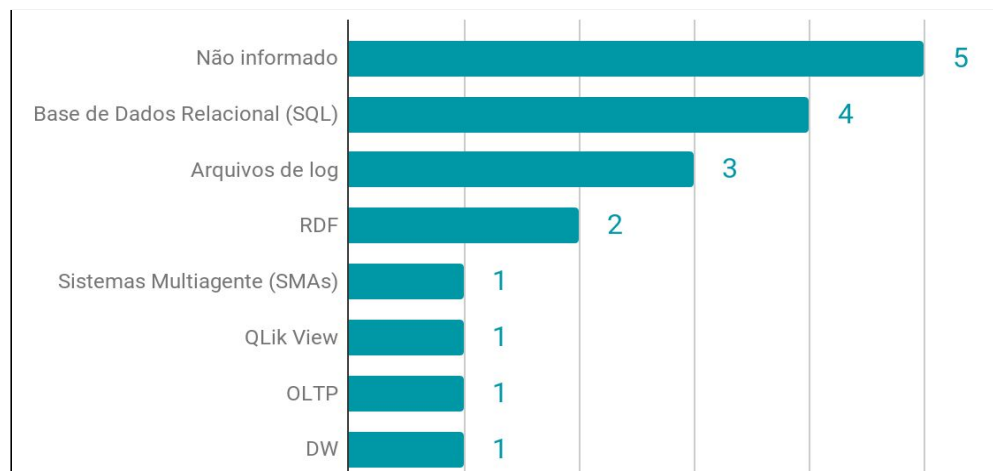


Figura 4. Gráfico referente às abordagens ou tecnologias utilizadas na coleta de dados

QP5: Que abordagens/tecnologias de análise são as mais aplicadas?

A grande maioria dos trabalhos (64%) relatou utilizar técnicas de mineração de dados para análise de aprendizagem. Conforme ilustrado na Figura 5, as tecnologias de análise mais utilizadas foram: Algoritmo de Classificação K-Means (utilizado por 23% dos estudos); Algoritmo para criação de árvore de decisão J48 (utilizado por 23% dos estudos); Métodos Estatísticos entre eles R, R Square, Adjusted Square (utilizado por 23% dos estudos).

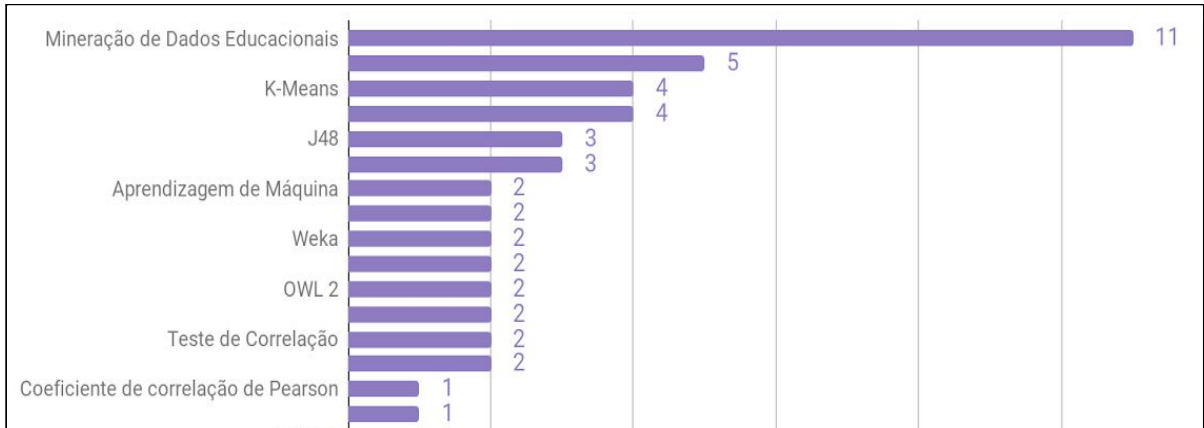


Figura 5. Gráfico referente às abordagens ou tecnologias utilizadas na análise dos dados

Percebe-se que diversas outras técnicas e ferramentas de análise de dados foram relatadas nos estudos, entre elas: Aprendizado de Máquina, Random Forest, Sistema Weka, Teste de Correlação, Coeficiente de correlação de Pearson, Verificação de Outliers, Algoritmo Z-Score, Algoritmo ID3, Algoritmos de Clusterização Hierárquica, Desvio padrão, Redes Bayesianas, Algoritmo de Simple Logistic, Algoritmo Multilayer Perceptron, Sistema Qlik View, Sistema Moodle Predicta, Sistema Pentaho, Cubo OLAP, Índice de Silhueta, Biblioteca Javascript jsChart, Linguagem SPARQL, Algoritmos Fuzzy, Técnicas de Regressão, Teste Spearman, Ferramenta Protege, Ferramenta MultiTrail, Ferramenta CMap Tools, Algoritmos de Processamento de Linguagem Natural, Regras XML, Sistema CoGrOO, Algoritmo Expectation-Maximization e Sistema FAG.

QP6: Os estudos apresentam resultados positivos as abordagens de análise de aprendizagem propostas?

Na grande maioria dos estudos (94%) os autores consideraram positivos os resultados do seu trabalho. Somente em 1 trabalho os autores não diagnosticaram o resultado. Verificamos que o respectivo trabalho não propõe estudo de caso. Acreditamos, desta forma, que isso tenha dificultado o diagnóstico dos resultados do estudo.

QP7: Que ambientes virtuais de aprendizagem se tem utilizado de análise de aprendizagem?

Nos trabalhos selecionados, a utilização do ambiente virtual de aprendizagem Moodle predomina em 70% das pesquisas. Acredita-se que esse número elevado se dá pelo Moodle ser uma plataforma baseada em software livre e já ser consagrado no meio acadêmico. Porém,

ainda foi levantado o uso do PBLMaestro (Id: 26878), um software desenvolvido durante a pesquisa dos autores. Outros cinco estudos (Ids: 26956, 26961, 26973, 27094) não informaram que ambiente virtual de aprendizagem foi utilizado.

5. Considerações Finais

No Brasil, estudos que trabalham com a aplicação de análise de aprendizagem em ambientes virtuais de ensino ainda são escassos. O que demonstra ainda a carência de evidências que comprovem como LA pode ser bem-sucedida quando aplicada em ambientes de ensino brasileiros. Esta pesquisa buscou identificar como a análise de aprendizagem tem sido aplicada em ambientes virtuais de aprendizagem utilizado por instituições de pesquisa/ensino nacionais. Verificamos que, embora LA possa ser aplicada em qualquer modalidade de ensino, a maioria dos trabalhos brasileiros utilizam-se de LA em cursos à distância. Além disso, verificamos que grande parte dos trabalhos aplicam essa tecnologia com objetivo de verificar os índices de performance e riscos de evasão dos alunos. Acreditamos que há altas expectativas para esse campo de pesquisa que ainda não foram atingidas e/ou divulgadas. Com o uso de LA é possível identificar, por exemplo, os padrões comportamentais e sociais que impactam no aprendizado do aluno. Esperamos através desta pesquisa promover ferramentas que utilizam LA e sintetizar os dados o estado atual deste campo de pesquisa a nível nacional. Além disso, apoiar a construção de novas ferramentas que auxiliem a tomada de decisão docente, no que diz respeito a realização de intervenções pedagógicas, durante o processo de ensino, seja ele presencial ou à distância.

6. Referências

ALMAZROUI, Yousef A. **A survey of Data mining in the context of E-learning.** International Journal of Information Technology & Computer Science (IJITCS), v. 7, n. 3, p. 8-18, 2013.

BAKER, R. S. J. D.; YACEF, K. **The state of educational data mining in 2009: a review and future visions.** Journal of Educational Data Mining, v. 1, n. 1, p. 3-17, 2009.

BAKER, R.; SIEMENS, G. Educational data mining and learning analytics. In: SAWYER, K. (Ed.). **The Cambridge handbook of the learning sciences**. 2. ed. Cambridge: Cambridge University Press, 2014. p. 253-274.

BAYER, J. et al. Predicting drop-out from social behaviour of students. In: **INTERNATIONAL CONFERENCE ON EDUCATIONAL DATA MINING, 5.**, 2012, Chania. Anais. 2012.

BIOLCHINI, Jorge et al. **Systematic review in software engineering**. System Engineering and Computer Science Department COPPE/UFRJ, Technical Report ES, v. 679, n. 05, p. 45, 2005.

BLIKSTEIN, Paulo. Using learning analytics to assess students' behavior in open-ended programming tasks. In: **Proceedings of the 1st international conference on learning analytics and knowledge**. ACM, 2011. p. 110-116.

DAWSON, Shane Peter et al. **Learning or performance: Predicting drivers of student motivation**. 2009.

DRINGUS, Laurie, P., ELLIS, Timothy. (2005). **Using data mining as a strategy for assessing asynchronous discussion forums**. Computer and Education Journal, 45, 141–160.

EISEVIER. **Scopus**. Disponível em: <https://service.elsevier.com/app/answers/detail/a_id/15534/supporthub/scopus/#tips>.

Acesso em: 21 jul. 2018.

FERGUSON, Rebecca et al. **Research evidence on the use of learning analytics: Implications for education policy**. 2016.

KITCHENHAM, B. A; CHARTERS, S. (2007) **Guidelines for performing systematic literature reviews in software engineering - version 2.3**, Keele/Staffs-UK and Durham-UK.

KOTSIANTIS, S. B. **Use of machine learning techniques for educational proposes: a decision support system for forecasting students' grades**. Artificial Intelligence Review v. 37, n. 4, p.331-344, 1 Apr. 2012

MÁRQUEZ-VERA, C. et al. **Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data.** Applied Intelligence, v. 38, n. 3, p. 315-330, 1 Apr. 2013.

ROMERO, C., López, M.-I., Luna, J.-M., and Ventura, S. (2013). **Predicting students' final performance from participation in on-line discussion forums.** Computers & Education, 68:458 – 472.

SAKOWSK, Patrícia Alessandra Moritai; TÓVOLI, Marina Haddad. **Perspectivas da Complexidade para a Educação no Brasil.** Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA), 2015.