

# O Uso de Learning Analytics em Ambientes de Aprendizagem Online: um Mapeamento Sistemático da Literatura

*Title: The Use of Learning Analytics in Online Learning Environments: a Systematic Literature Mapping*

*Michael Miller Rodrigues Cardoso*  
*Universidade Federal de Alagoas*  
*mmrc@ic.ufal.br*

*João Victor Falcão Santos Lima*  
*Universidade Federal de Alagoas*  
*jvysl@ic.ufal.br*

*Márcio Henrique Vieira de Oliveira*  
*Universidade Federal de Alagoas*  
*mhvo@ic.ufal.br*

*Ranilson Oscar Araujo Paiva*  
*Universidade Federal de Alagoas*  
*ranilsonpaiva@ic.ufal.br*

## Resumo

Nos últimos anos houve um grande aumento no número de pesquisas e possibilidades de aplicação para a área de Learning Analytics, no que diz respeito à medição, coleta, análise e comunicação de dados dos alunos e seus contextos na educação. Este estudo realiza uma avaliação do estado atual do campo de Learning Analytics por meio da análise de artigos científicos que abordam o uso em ambientes educacionais de aprendizagem online. Seguindo um planejamento/protocolo e condução com auxílio da plataforma Sumarize, foi realizado um mapeamento sistemático que resultou em 38 estudos finais que focam no objetivo de entender como Learning Analytics é utilizada para ajudar a identificar os estudantes que estão enfrentando dificuldades em cursos online. Os estudos foram analisados detalhadamente com base no objetivo principal mencionado anteriormente e nos demais objetivos de pesquisa definidos. Os resultados forneceram uma visão geral da área de Learning Analytics e dos diferentes problemas que podem ser solucionados através de seu uso, destacando a identificação de estudantes com dificuldades, na tentativa de solucionar o problema da reprovação e abandono na aprendizagem a distância.

**Palavras-chave:** *Análise de Aprendizagem; Ambientes de Aprendizagem Online; Visualização de Dados; Previsão de Estudantes em Risco; Educação Online; Mapeamento Sistemático da Literatura.*

## Abstract

In recent years, there has been a great increase in the number of researches and application possibilities for an area of Learning Analytics, with regard to the measurement, collection, analysis and communication of data from students and their contexts in education. This study assesses the current state of the field of learning analytics through the analysis of scientific articles that address its use in online learning educational environments. Following a planning/protocol and conducting with the help of the Sumarize platform, a systematic mapping was carried out that resulted in 38 final studies that focus on the objective of understanding how Learning Analytics is used to help identify students who are facing difficulties in online courses. The studies were analyzed in detail based on the main objective mentioned above and the other defined research objectives. The results provided an overview of the area of Learning Analytics and the different problems that can be solved through its use, highlighting the identification of students with difficulties, in an attempt to solve the problem of failure and dropout in distance learning.

**Keywords:** *Learning Analytics; Online Learning Environments; Data Visualization; Prediction of Students at Risk; Online education; Systematic Literature Mapping.*

## 1. Introdução

Learning Analytics (LA) é a medição, coleta, análise e comunicação de dados sobre os alunos e os seus contextos (Gašević et al., 2015), com o propósito de compreender e otimizar a aprendizagem e os ambientes online em que ocorre, conforme definido em 2011 no primeiro International Learning Analytics and Knowledge Conference (LAK), esta definição geral ainda é válida mesmo com o crescimento do campo.

Nos últimos anos houve um grande aumento no número de pessoas que pesquisam o aprendizado e o ensino, acompanham o progresso dos alunos, analisam dados escolares ou universitários, elaboram avaliações e usam evidências para melhorar o ensino e aprendizagem (Dawson et al., 2014) (Siemens, 2013). Com o objetivo de identificar o surgimento de tendências na área, os problemas que podem ser solucionados com a aplicação de LA e os principais métodos utilizados para implantar análises em ambientes de aprendizagem online, foi realizado um mapeamento sistemático das pesquisas existentes a fim de verificar a natureza, a extensão e a quantidade de estudos publicados nesta área.

O processo de Learning Analytics deve ser usado para medir e coletar dados sobre os alunos e a aprendizagem, com o objetivo de melhorar o ensino e a prática de aprendizagem por meio da análise dos dados. A partir da coleta dos dados, algoritmos/técnicas de LA são utilizadas para analisar o desempenho do aluno e fazer previsões de risco de reprovação e/ou abandono. Por fim, os professores podem visualizar estas análises e previsões através de gráficos e/ou tabelas e realizar intervenções para auxiliar os estudantes que estão enfrentando dificuldades no ambiente de aprendizagem online (Clow, 2012).

Historicamente, alguns dos usos mais comuns de Learning Analytics são a previsão do sucesso acadêmico do aluno e, mais especificamente, a identificação de alunos que correm o risco de serem reprovados em um curso ou abandonar os estudos (Gašević et al., 2016) (Leeuwen et al., 2019). No geral, o uso de Learning Analytics pode ser muito útil em ambientes educacionais para aplicação das seguintes etapas: monitoramento dos alunos, análise dos dados, previsão de estudantes em risco, intervenções por parte dos professores, tutoria, avaliação, adaptação, personalização, recomendação, feedback, conscientização e reflexão.

Além de LA, existe outro processo muito popular para análise de dados, denominado Educational Data Mining (EDM). Os dois campos de pesquisa são relativamente recentes e cada vez mais populares no ramo da coleta, análise e interpretação de dados educacionais. A grande diferença entre ambos é a seguinte: a EDM possui foco nos dados e nas técnicas/algoritmos utilizados; enquanto Learning Analytics está relacionada aos aspectos educacionais além da aprendizagem, portanto, enfoca como o cenário educacional (professores e alunos) pode se beneficiar do uso dessas técnicas (Gašević et al., 2015).

Para uma melhor interpretação da comparação entre LA e EDM, podemos citar Siemens e Baker (2012), que listaram as principais diferenças, entre elas:

- Em LA a abordagem é holística (trata do sistema como um todo) e em EDM é reducionista (divide o problema em componentes que são tratados individualmente);
- O foco na adaptação automática na mineração de dados educacionais, enquanto Learning Analytics foca em informar e capacitar professores e alunos;
- Diferenças nas técnicas e algoritmos utilizados: enquanto Learning Analytics considera Social Network Analysis (SNA), Sentiment Analysis, entre outros, a mineração de dados educacionais faz mais uso de técnicas de machine learning, como classificação, clusterização, modelagem Bayesiana, entre outros, que em muitos casos também são utilizados em LA.

Resumidamente, Learning Analytics engloba o uso de vários métodos, que incluem técnicas de análise de dados visuais, SNA, análises estatísticas, além de algoritmos provenientes da mineração de dados como técnicas de predição, clusterização, correlações, entre outros, incluindo a separação de dados valiosos para auxiliar professores no processo de ensino-aprendizagem (Avella et al, 2016). Learning Analytics aumenta a conscientização dos alunos e professores em suas situações atuais, podendo ajudá-los a tomar decisões construtivas e executar de forma mais eficaz suas tarefas no ambiente educacional online (Scheffel et al., 2014).

A escolha de conduzir um mapeamento sistemático para entender o campo de Learning Analytics se deve pelo fato de que esse modelo de pesquisa fornece uma estrutura dos tipos de relatórios de pesquisa e resultados que foram publicados, categorizando-os e fornecendo um resumo visual e/ou descritivo dos dados, a partir da sumarização de seus resultados. Devido à existência de diferentes tecnologias e metodologias para o desenvolvimento de LA, além dos desafios enfrentados por esta área, o mapeamento possui foco na identificação de estudos que abordam a aplicação prática de Learning Analytics em ambientes de aprendizagem online. Além disso, os critérios de seleção e avaliação dos artigos têm como alvo estudos com o objetivo de auxiliar os professores na identificação de estudantes enfrentando dificuldades em um determinado curso.

Inicialmente, foi elaborado o protocolo de busca e seleção, em seguida foram coletados 3315 artigos, que primeiro foram analisados por título, resumo e palavras-chave e depois analisados por completo, restando um total de 38 artigos. A partir da sumarização dos resultados do mapeamento, os principais objetivos do uso de Learning Analytics e as tecnologias e algoritmos mais utilizados em ambientes de aprendizagem online foram identificados e foi possível destrinchar os diferentes problemas da aprendizagem a distância que podem ser solucionados através da sua aplicação, entre eles, identificar estudantes em risco em ambientes online, definida como a principal questão de pesquisa.

Na seção 2 temos a metodologia, vindo em seguida os resultados do mapeamento (seção 3), as discussões acerca dos resultados obtidos (seção 4), e por fim, as conclusões e considerações finais (seção 5) e as referências bibliográficas.

## 2. Metodologia

O mapeamento sistemático da literatura foi desenvolvido com auxílio da plataforma Sumarize<sup>1</sup> e conduzido com base no processo de execução descrito por Kitchenham e Charters (2007), que descreve três fases principais: planejamento, condução e relatório, conforme ilustra a Figura 1. Este processo serve tanto para um mapeamento sistemático como para uma revisão sistemática da literatura (RSL). Ele foi escolhido devido ao fato de ser fundamentado em outros protocolos amplamente utilizados na pesquisa médica baseada em evidência, além de ser bastante utilizado tanto na área de computação, em geral, quanto nos mapeamentos e revisões sistemáticas da área de Informática na Educação.

A plataforma Sumarize foi utilizada nas seguintes etapas: (1) Importação dos metadados dos artigos retornados nas buscas, gerando uma lista de artigos selecionados; (2) Primeira seleção dos artigos com base nos títulos abstract e palavras-chave, marcando como aceito (associando a aceitação a um critério de inclusão) ou rejeitado (associando a rejeição a um critério de exclusão); (3) Segunda seleção dos artigos com base na leitura por completo, marcando como aceito ou rejeitado; (4) Condução da avaliação de qualidade, marcando a pontuação dos artigos em cada critério definido, obtendo assim, a nota final dos mesmos (a própria plataforma marcava quais

---

<sup>1</sup> Plataforma Sumarize: <https://sumarize.evidencias.nees.com.br/>

artigos foram aprovados e reprovados na avaliação de qualidade, a partir do limiar mínimo definido).

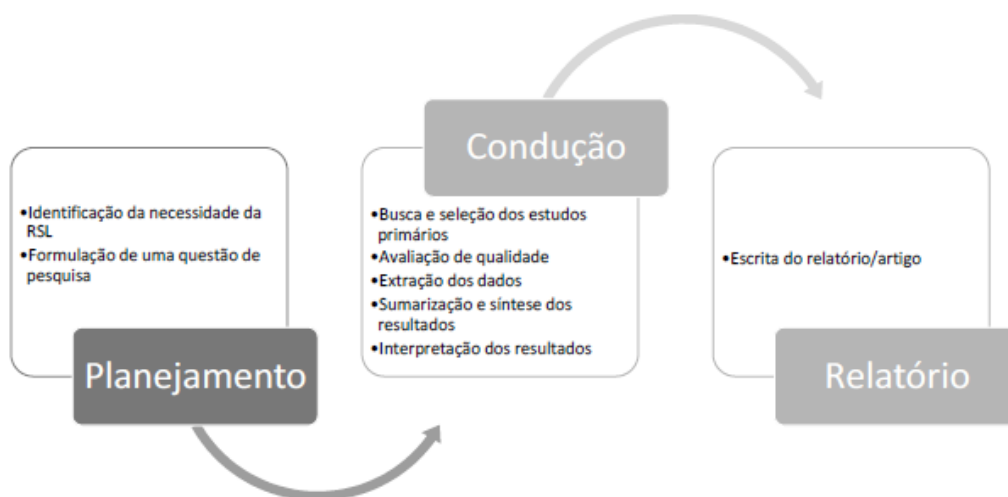


Figura 1: Processo de execução de um mapeamento, descrito por Kitchenham e Charters.

Inicialmente, na fase de planejamento, foi definido um protocolo de pesquisa com os seguintes tópicos: formulação das questões de pesquisa (RQs - Research Questions), palavras-chaves e sinônimos, string de busca, fontes (bibliotecas digitais), critérios de inclusão e de exclusão. De acordo com os objetivos definidos para a condução do mapeamento, foram formuladas as seguintes questões de pesquisa:

- RQ0: De que forma Learning Analytics é utilizada para ajudar a identificar os estudantes que estão enfrentando dificuldades em um curso online, considerando cada paper/estudo/experimento analisado?
- RQ1: Qual o propósito/objetivo do uso de LA em cada paper/estudo/experimento analisado?
- RQ2: Quais as técnicas/algoritmos/ferramentas utilizadas para cada propósito de uso de Learning Analytics em cada paper/estudo/experimento analisado?
- RQ3: Quais os sistemas/softwarees utilizados em cada paper/estudo/experimento analisado?
- RQ4: Quais dados foram usados para cada uma dessas técnicas de Learning Analytics em cada paper/estudo/experimento analisado?
- RQ5: Como os dados foram coletados, pré-processados em cada paper/estudo/experimento analisado?
- RQ6: Quais as formas de visualização dos dados dos estudantes utilizadas em cada paper/estudo/experimento analisado?
- RQ7: Durante o processo, houve a participação colaborativa das inteligências humana e artificial?
- RQ8: Quais os benefícios/melhorias que foram promovidos pelo uso de Learning Analytics em cada paper/estudo/experimento analisado?

- RQ9: A quem se destinava, diretamente, o resultado da aplicação de LA em cada paper/estudo/experimento analisado?
- RQ10: Qual a percepção dos usuários finais em relação ao uso de Learning Analytics em cada paper/estudo/experimento analisado?

Após a formulação das questões de pesquisa, os demais tópicos do protocolo foram elaborados, desde a definição da string de busca a partir das palavras-chave, a elaboração dos critérios de inclusão e exclusão e a estratégia de busca dos artigos, para finalizar a fase de planejamento, como é descrito a seguir.

Para auxiliar na definição da string de busca e nas questões que foram exploradas no presente estudo, foi utilizado o framework PICOC, também citado por Kitchenham e Charters, que estabelece 5 termos: População, Intervenção, Comparação, Outcome/Resultado e Contexto. Portanto, para nossa pesquisa, foram definidos os seguintes termos:

**P** (População) - Professores e estudantes que utilizam ambientes online de aprendizagem;

**I** (Intervenção) - Utilização das técnicas de Learning Analytics;

**C** (Comparação) - Não se aplica, pois o estudo não demanda a realização de comparações com outras tecnologias da área;

**O** (Outcome/Resultado) - Identificação e acompanhamento de estudantes com dificuldades em cursos online;

**C** (Contexto) - Ambientes online de aprendizagem.

A partir da estruturação das questões de pesquisa em palavras-chave logicamente organizadas e do framework PICOC, foi elaborada a seguinte string de busca:

*("learning analytics" OR "e-learning analytics" OR "learning analysis" OR "pedagogical analysis") AND ("online learning" OR "distance education" OR "distance learning" OR "e-learning" OR "online education" OR "virtual education" OR "virtual learning" OR "web-based education" OR "web-based learning") AND ("environment\*" OR "approach" OR "approaches" OR "framework\*" OR "technique\*" OR "tool\*")*

Vale ressaltar que o processo para definir os sinônimos da string foi feito manualmente através de testes de busca no Google Scholar e nas bibliotecas digitais escolhidas (**Tabela 2**), identificando através dos resultados obtidos, quais termos deveriam ser incluídos.

Uma vez que a string de busca foi estruturada e validada, foram definidos os critérios de inclusão e de exclusão para seleção dos artigos (**Tabela 1**):

Tabela 1: Critérios de inclusão e exclusão.

Critérios de inclusão	Critérios de exclusão
Publicado em Inglês	Estudos não publicados em Inglês
Estudos com foco no uso de learning analytics em ambientes online	Estudos duplicados
Estudos focados em ambientes online/híbridos de aprendizagem	Literatura cinza (livros)
Estudos que abordem learning analytics	Estudos que não abordem learning analytics
Estudos que fornecem respostas para as perguntas da pesquisa	Estudos não relacionados às questões de pesquisa

Estudos que utilizem tecnologias/ferramentas	Outras revisões/mapeamentos sistemáticos
Fontes primárias	Fontes secundárias
Publicado desde 2010	Estudos não publicados entre 2010-2020
Artigos completos	Short-papers
-----	Não foi encontrado o artigo completo

O processo para avaliar os critérios de inclusão e exclusão foi dado através da investigação em outros mapeamentos sistemáticos e do método proposto por Kitchenham e Charters (2007), levando também em conta as questões de pesquisa do presente estudo, para selecionar somente as publicações relevantes. Vale ressaltar que estes critérios foram definidos em conjunto pelos pesquisadores envolvidos no mapeamento, assim como as questões da avaliação de qualidade.

O presente estudo teve sua condução finalizada na primeira metade de 2021, por conta disso, o intervalo de busca foi até o ano de 2020 (até o fim do ano anterior). Em um trabalho futuro, o período será estendido, avaliando o uso de Learning Analytics pós-pandemia.

Em relação à exclusão de estudos não publicados em inglês, foi definido como critério de seleção considerar somente a língua científica internacional (atual), para que pesquisadores de qualquer nacionalidade, com suas pesquisas publicadas em inglês, fossem incluídos no mapeamento. Com isso, é possível trazer um panorama internacional da área.

Com a string de busca e os critérios de exclusão e inclusão definidos, o último passo foi estabelecer a estratégia de busca dos artigos. A estratégia definida foi recuperar artigos retornados de bibliotecas digitais a partir da string de busca, e, em seguida, conduzir a seleção e análise destes a partir dos critérios de inclusão e exclusão (partindo da fase de planejamento para condução).

Foram escolhidas as principais bibliotecas digitais na área da computação (**Tabela 2**). Outras bibliotecas indexam os mesmos artigos.

Tabela 2: Bibliotecas digitais selecionadas.

<b>Fonte</b>	<b>Endereço</b>
ACM Digital Library	<a href="http://portal.acm.org/">http://portal.acm.org/</a>
Ei Compendex	<a href="https://www.elsevier.com/solutions/engineering-village">https://www.elsevier.com/solutions/engineering-village</a>
IEEE Digital Library	<a href="http://ieeexplore.ieee.org/">http://ieeexplore.ieee.org/</a>
ISI Web of Science	<a href="http://www.isiknowledge.com/">http://www.isiknowledge.com/</a>
Science@Direct	<a href="http://www.sciencedirect.com/">http://www.sciencedirect.com/</a>
Scopus	<a href="http://www.scopus.com/">http://www.scopus.com/</a>
Springer Link	<a href="http://link.springer.com/">http://link.springer.com/</a>

Após os passos anteriores, foram definidas as questões da avaliação de qualidade, com o objetivo de aprovar para o mapeamento sistemático somente artigos relevantes da área de

pesquisa. Foi definida uma pontuação de 0 a 2 para cada questão da avaliação de qualidade, exceto para as perguntas sobre o tipo e o ano de publicação (**Tabela 3**). Para que um artigo fosse aprovado, deveria alcançar um limiar mínimo de 13 pontos de um total de 19, somando as pontuações em cada critério de qualidade (pergunta):

Tabela 3: Critérios de qualidade.

Pergunta	Pontuações possíveis
O artigo foca no uso de LA pelos professores no apoio aos estudantes?	Sim (2); Parcialmente (1); Não (0).
Qual o fator de impacto da publicação?	Maior que 4 (2); Maior que 3, até 4 (1,75); Maior que 2.5, até 3 (1,5); Maior que 2, até 2.5 (1,25); Maior que 1.5, até 2 (1); Maior que 1, até 1.5 (0,75); Maior que 0.5, até 1 (0,5); Até 0.5 (0,25); Não encontrado (0).
Qual o número de citações do artigo?	Maior que 75 (2); 50 a 75 (1,5); 25 a 49 (1); 1 a 24 (0,5); 0 (0).
O artigo é recente? (a partir de 2016)	A partir de 2016 (1); Entre 2010 e 2015 (0).
Qual o tipo de publicação? (conferência ou periódico)	Periódico (2); Conferência (1).
Realiza estudo empírico?	Sim, realiza experimento controlado (2); Sim, realiza estudo menos rígido (1); Não (0).
O artigo segue algum processo? (sequência de passos organizada em etapas para aplicação e/ou investigação do uso de LA)	Sim, e é validado (2); Sim, mas não é validado (1); Não (0).
O artigo demonstra aplicações práticas de Learning Analytics?	Sim (2); Parcialmente (1); Não (0).
O artigo apresenta a implementação da tecnologia abordada?	Implementa protótipo / utiliza software existente (2); Não implementa, mas passa arquitetura / requisitos (1); Não aborda implementação (0).
O artigo aborda a previsão de estudantes em risco de abandono?	Sim (2); Parcialmente (1); Não (0).

Vale ressaltar que o FI (Fator de Impacto) é importante para selecionar as publicações mais relevantes, publicadas nas conferências e/ou periódicos mais populares e com maior número de citações na área de Learning Analytics. Além disso, destacamos que a pontuação das questões da avaliação de qualidade foi definida de acordo com a quantidade de respostas em cada pergunta.

A última etapa antes de partir para a condução do mapeamento foi elaborar um formulário de extração dos dados sob medida com base nas perguntas de pesquisa, para obter todas as informações relevantes de cada um dos estudos incluídos no mapeamento, permitindo realizar uma análise geral e também a sumarização dos dados. O formulário inclui as seguintes perguntas: (1) Identificador do estudo; (2) Fonte e tipo de fonte; (3) Ano de publicação; (4) Autores e afiliação dos autores; (5) Título e palavras-chave; (6) Número de citações; (7) Data de extração dos dados; (8) Problema e Justificativa; (9) Objetivo(s) do uso de LA; (10) Proposta; (11) Propõe um sistema/ferramenta?; (12) Técnicas e/ou algoritmos utilizados; (13) Software utilizado (se existir); (14) Dados utilizados e como foram coletados; (15) Método(s) utilizado(s); (16) Formas de visualização utilizadas; (17) Houve colaboração entre inteligências humana e artificial?; (18) Realiza estudo empírico?; (19) Tipo de estudo; (20) População; (21) Percepções dos usuários.

Após a fase de planejamento (criação de um protocolo para o mapeamento sistemático), foi dado início à fase de condução, onde foi realizada a busca e seleção dos artigos, avaliação de qualidade com base em critérios estabelecidos e a extração de dados dos estudos relevantes. No fluxograma (**Figura 2**), podemos visualizar todo o processo de seleção dos artigos:

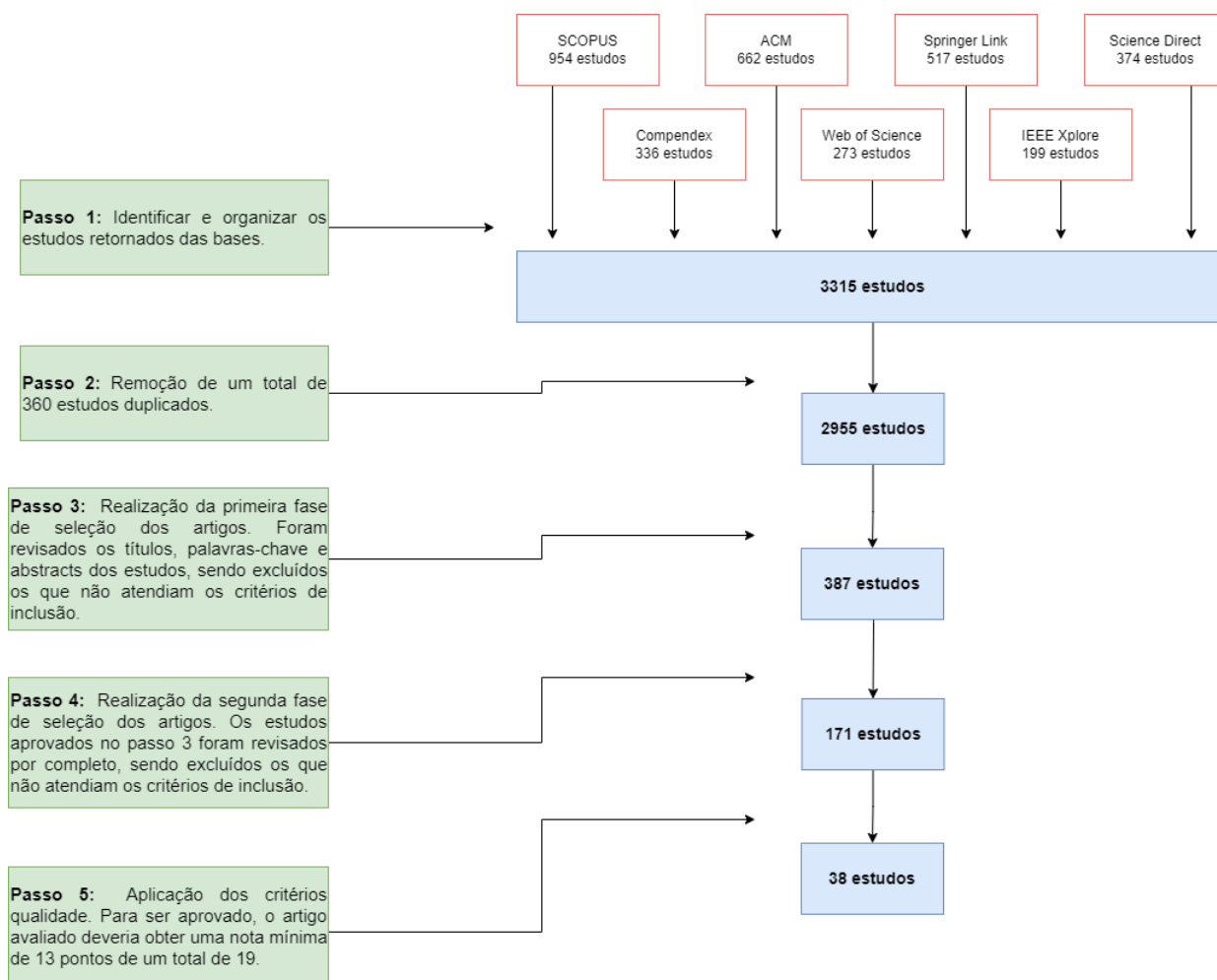


Figura 2: Fluxograma do processo de seleção de artigos.



Inicialmente, foram coletados 3315 estudos a partir da busca nas fontes digitais escolhidas, nos quais foram removidos 360 estudos duplicados e após a primeira e segunda fase de seleção dos artigos restaram somente 171 estudos (**Tabela 4**). Durante a primeira fase de seleção, os artigos foram analisados quanto aos critérios de inclusão e exclusão, através da leitura dos títulos, palavras-chave e resumo, enquanto na segunda fase os estudos foram lidos por completo.

Tabela 4: Quantidade de artigos por fonte após a segunda fase de seleção.

Fonte	Quantidade	%
Scopus	39	22.81
ACM	38	22.22
SpringerLink	19	11.11
ScienceDirect	20	11.70
Web of Science	34	19.88
IEEE Xplore	20	11.70
Ei Compendex	1	0.58
Total	171	100.0

Em seguida, os 171 artigos selecionados foram avaliados com base nos critérios de qualidade estabelecidos previamente no protocolo (**Tabela 5**), onde somente os artigos com um limiar de qualidade superior a 68.4% foram aprovados, restando um total de 38 estudos finais (**Tabela 6**). Foi definido um limiar de 13 pontos devido ao fato de ser o valor inteiro que torna a nota mínima para aprovação de um artigo mais próxima da proporção 7/10. Vale ressaltar que, apesar da pontuação máxima ter sido de 19 pontos, nenhum artigo obteve nota maior que 16.

Tabela 5: Notas dos artigos nos critérios de qualidade.

Notas	Quantidade	%
(6, 8]	15	8.77
(8, 10]	49	28.65
(10, 13)	69	40.35
[13, 15]	28	16.37
(15, 16]	10	5.84
(16, 19]	0	0.0
Total	171	100.0

Tabela 6: Artigos finais por fonte

Notas	Quantidade	%
Scopus	4	10.53
ACM	2	5.26
SpringerLink	7	18.42
ScienceDirect	11	28.95
Web of Science	13	34.21
IEEE Xplore	1	2.63
Ei Compendex	0	0.0
Total	38	100.0

### 3. Resultados

Após a conclusão do planejamento do mapeamento sistemático e das fases de condução que foram apresentadas na metodologia deste estudo, os resultados obtidos foram sumarizados e geradas diversas visualizações, com o objetivo de fornecer auxílio para responder às questões de pesquisa elaboradas no protocolo.

Durante a presente seção, são apresentados os itens com maiores quantidades/porcentagens e somente alguns autores são mencionados para cada resultado das perguntas de pesquisa, especificamente os 3 autores com maior número de citações (H-index). Caso deseje visualizar os estudos de todos os autores com seus respectivos resultados, acesse a tabela disponibilizada na nota de rodapé<sup>2</sup>.

- RQ1: Qual o propósito/objetivo do uso de LA no paper/estudo/experimento?

Observando a **Tabela 7**, temos que o principal objetivo do uso de Learning Analytics é a previsão de estudantes em risco em um curso/disciplina, abordado em 26 dos 38 artigos (68%), com destaque para os estudos (Gašević et al., 2016) e (Xing et al., 2016), com maior número de citações. Em seguida, temos a visualização de informações presente em 13 artigos (34%), a personalização e adaptação em 12 estudos (31%), o foco em promover a autorregulação dos estudantes em 8 artigos (21%) e por fim, com menor frequência, a aprendizagem cognitiva em 3 dos 38 artigos (8%), respondendo então, a questão de pesquisa RQ1.

<sup>2</sup> Tabela de resultados: <https://tinyurl.com/ArticlesLA>

Tabela 7: Objetivos do uso de Learning Analytics.

Objetivos	Quantidade	%	Autores/Ano
Previsão de estudantes em risco	26	68	Dragan Gasevic, 2016; Wanli Xing, 2016; Christothea Herodotou, 2019.
Visualização de informações	13	34	J.A. Ruipérez-Valiente, 2017; Jalal S. Alowibdi, 2019; A. van Leeuwen, 2019.
Personalização e adaptação	12	31	Ahmed M. F. Yousef, 2016; Abelardo Pardo, 2017; Fredys A. S. Herrera, 2019.
Promover a autorregulação	8	21	Dragan Gašević, 2016; Ji Won You, 2016; Quan Nguyen, 2019.
Aprendizagem cognitiva	3	8	Abelardo Pardo, 2017; Gaoxia Zhu, 2019; Jalal S. Alowibdi, 2019.

- RQ2: Quais as técnicas/algoritmos/ferramentas utilizadas para cada propósito de uso de Learning Analytics no paper/estudo/experimento?

A partir da visualização da **Tabela 8**, é possível responder a questão de pesquisa RQ2, concluindo que as técnicas e/ou algoritmos mais utilizados nos artigos do mapeamento foram as regressões, que estavam presentes em 13 estudos (34%), vindo em seguida a análise de variância, em 8 dos 38 estudos (21%), análise de cluster (16%) e correlação (16%), Social Network Analysis em 5 dos 38 artigos (13%), Random Forest (13%), outras técnicas/algoritmos listados na nota de rodapé<sup>3</sup> (13%), Naive Bayes em 4 artigos (10%) e por fim, os testes de hipótese, também presentes em 4 dos 38 artigos (10%).

Tabela 8: Técnicas/algoritmos mais utilizados.

Técnica/algoritmo	Quantidade	%	Autores/Ano
Regressões	13	34	Dragan Gasevic, 2016; Nick Z. Zacharis, 2015; Quan Nguyen, 2019.
Análise de Variância	8	21	Dragan Gasevic, 2016; Ji Won You, 2016; Nick Z. Zacharis, 2015.
Análise de Cluster	6	16	Shu-Fen Tseng, 2016; Viet Anh Nguyen, 2018; Nick Z. Zacharis, 2015.
Análise de Correlação	6	16	Dragan Gasevic, 2016; Abelardo Pardo, 2016; Uno Fors, 2017.

<sup>3</sup> Outros algoritmos/técnicas: Árvores de decisão, redes neurais recorrentes (LSTM), Discourse Analysis, Sentiment Analysis e Principal Component Analysis (PCA).

SNA	5	13	A. Hernandez-Garcia, 2015; Stefan Trausan-Matu, 2015; Jorge Castellanos, 2017.
Random Forest	5	13	Viet Anh Nguyen, 2018; Gokhan Akcapinar, 2019; Du Q. Huynh, 2020.
Outros	5	13	Ahmed M. F. Yousef, 2016; Jalal S. Alowibdi, 2019; A. van Leeuwen, 2019.
Naive Bayes	4	10	Viet Anh Nguyen, 2018; Gokhan Akcapinar, 2019; Du Q. Huynh, 2020.
Testes de hipótese	4	10	Dragan Gasevic, 2016; Ji Won You, 2016; Nick Z. Zacharis, 2015.

- RQ3: Quais os sistemas/softwarewares mais utilizados nos papers/estudos/experimentos?

Quanto aos sistemas mais utilizados para aplicar o uso de Learning Analytics (questão de pesquisa RQ3), conforme a **Tabela 9**, temos que a maior parte dos artigos desenvolveram um ambiente próprio (29%) ou fizeram uso do Moodle (26%), com destaque para os artigos de Zacharis (2015) e Tseng et al. (2016), que estão entre os estudos com maior número de citações. Em seguida, temos outros sistemas/softwarewares<sup>4</sup> (sistemas utilizados em um único artigo, que foram colocados juntos para facilitar a visualização) presentes em 8 artigos (21%), depois temos a utilização da plataforma Blackboard (3 artigos / 8%), Open edX (3 artigos / 8%) e por fim, a plataforma Sowiso, utilizada em 2 dos estudos selecionados (5%).

Tabela 9: Sistemas/softwarewares utilizados nos estudos.

Sistema/software	Quantidade	%	Autores/Ano
Ambiente de aprendizagem online próprio	12	32	Dragan Gasevic, 2016; Shu-Fen Tseng, 2016; Á. Fidalgo-Blanco, 2015.
Moodle	10	26	Nick Z. Zacharis, 2015; Viet Anh Nguyen, 2018; Daniel A. Filvà, 2019.
Outros	8	21	Viet Anh Nguyen, 2018; Ji Won You, 2016; Abelardo Pardo, 2016.
Blackboard	3	8	Jalal S. Alowibdi, 2019; Quan Nguyen, 2018.
Open edX	3	8	Ruipérez-Valiente, 2017; Moreno-Marcos, 2018;

<sup>4</sup> Outros sistemas/softwarewares: Khan Academy, TICademia LMS, NextThought, MathTutor, OU Analyze, Elgg, Canvas LMS, Sakai LMS.

Sowiso	2	5	Quan Nguyen, 2019.
--------	---	---	--------------------

- RQ4: Quais dados foram usados para cada uma dessas técnicas de Learning Analytics no paper/estudo/experimento?

Durante o processo de condução do mapeamento sistemático já era possível perceber que a maioria dos artigos abordava o uso dos dados de interação dos estudantes no ambiente online. Esta hipótese é confirmada ao observar a **Tabela 10**, com a presença do uso dos dados de interações no ambiente online (número de cliques e acessos às páginas da plataforma) em 30 dos 38 artigos finais (79%), vindo em seguida as mensagens em discussões online, em 10 artigos (26%), os resultados de avaliações em 9 artigos (23%), percepções dos usuários (coleta de opiniões dos usuários em relação ao desempenho nos cursos online através de formulários) em 6 artigos (16%) e por fim, os dados demográficos dos estudantes em 3 estudos (8%), respondendo então, à questão de pesquisa RQ4.

Tabela 10: Dados utilizados nos estudos.

Dados	Quantidade	%	Autores/Ano
Interações no ambiente online	30	79	Nick Z. Zacharis, 2015; Dragan Gašević, 2016; Quan Nguyen, 2018;
Mensagens em discussões online	10	26	Ángel Fidalgo-Blanco, 2015; Wanli Xing, 2016; Gaoxia Zhu, 2019;
Resultados de avaliações	9	23	S. Akhtar, 2017; Daniel Burgos, 2019; Gokhan Akcapinar, 2019;
Percepções dos usuários	6	16	Ahmed Mohamed F. Yousef, 2016; J.A. Ruipérez-Valiente, 2017; Abelardo Pardo, 2017;
Dados demográficos	3	8	Reynol Junco, 2015; Christothea Herodotou, 2019; Hajra Waheed, 2020;

- RQ5: Como os dados foram coletados, pré-processados no paper/estudo/experimento?

Além de identificar os dados utilizados, também é de grande importância entender como estas informações são coletadas. Através da sumarização dos dados e da visualização da **Figura 3**, a questão de pesquisa RQ5 foi respondida, tendo como resultado que grande parte dos artigos finais (20 artigos, ou seja, 53%) abordavam a extração dos logs do sistema utilizado, ou seja, a coleta dos registros das interações realizadas no ambiente online. Em seguida, temos o uso de um plugin e/ou ferramenta proposta, presente em 11 artigos (29%), e por fim, com menor frequência temos a coleta de dados com questionário/entrevista que foi utilizado em 7 dos 38 artigos, ou seja, 18% dos estudos.

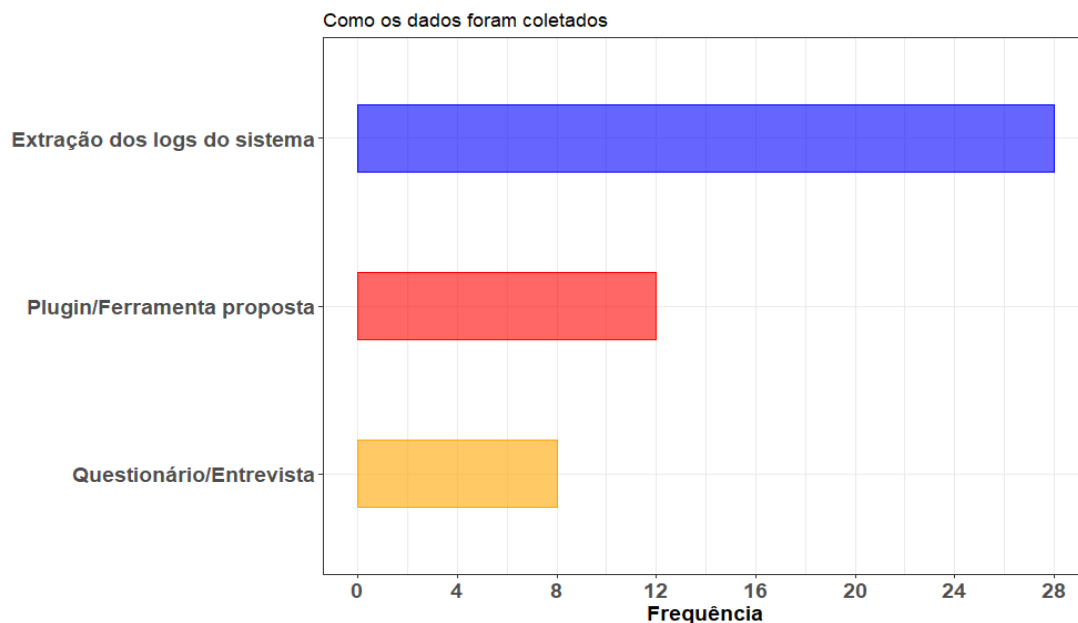


Figura 3: Como os dados foram coletados, pré-processados.

- RQ6: Quais as formas de visualização dos dados dos estudantes mais utilizadas no paper/estudo/experimento?

A partir da observação da Tabela 11, foi possível responder a questão de pesquisa RQ6, onde temos que as formas de visualização mais utilizadas são os gráficos estatísticos, presentes em 20 dos 38 estudos (52%), com destaque para o artigo de Xing et al. (2016), por ser um dos que apresentam mais visualizações desse tipo, vindo em seguida o uso de tabelas (12 artigos, 31%), de painéis de Learning Analytics e Social Network Analysis (11 estudos / 29%), com destaque para o artigo de Leeuwen et al. (2019), que foca no uso desses painéis, e por fim, com menor frequência, temos o uso de mapas de calor, presente em 4 artigos (11%).

Tabela 11: Formas de visualização mais utilizadas.

Tipo de visualização	Quantidade	%	Autores/Ano
Gráficos estatísticos	20	52	Angel Hernandez-Garcia, 2015; Nick Z. Zacharis, 2015; Wanli Xing, 2016;
Tabelas	12	31	Shu-Fen Tseng, 2016; Martin Hlostá, 2017; Gokhan Akcapinar, 2019;
Painéis de LA/SNA	11	29	J.A. Ruipérez-Valiente, 2017; Jalal S. Alowibdi, 2019; Anouschka van Leeuwen, 2019;
Mapa de calor	4	11	Christothea Herodotou, 2019; Hajra Waheed, 2020; Daniel Burgos, 2020;

- RQ7: Durante o processo, houve a participação colaborativa das inteligências humana e artificial?

Em 84% dos artigos finais (32 dos 38 estudos), houve a colaboração entre as inteligências humana e artificial, ou seja, a população dos estudos selecionados, seja professor ou aluno, interagiu com algum sistema ou ferramenta que aplica métodos inteligentes para auxiliá-lo, como previsão antecipada de estudantes em risco com análises estatísticas, aprendizado de máquina, entre outras técnicas. Somente em 16% dos estudos não houve colaboração entre a inteligência humana e artificial.

- RQ8: Quais os benefícios/melhorias que foram promovidos pelo uso de Learning Analytics?

Ao visualizar a **Figura 5**, temos os principais benefícios e melhorias que foram promovidos pelo uso de Learning Analytics com base nos artigos selecionados, onde podemos observar que o principal benefício foi a melhoria dos resultados acadêmicos, em 26 dos 38 artigos (68%), além da otimização do tempo dos professores no uso do ambiente de aprendizagem online (20 artigos / 52%), vindo em seguida o benefício de promover ensino personalizado (10 estudos / 26%) e o aumento da motivação dos estudantes, presente em 6 dos 38 artigos (16%), respondendo então, a questão de pesquisa RQ8.

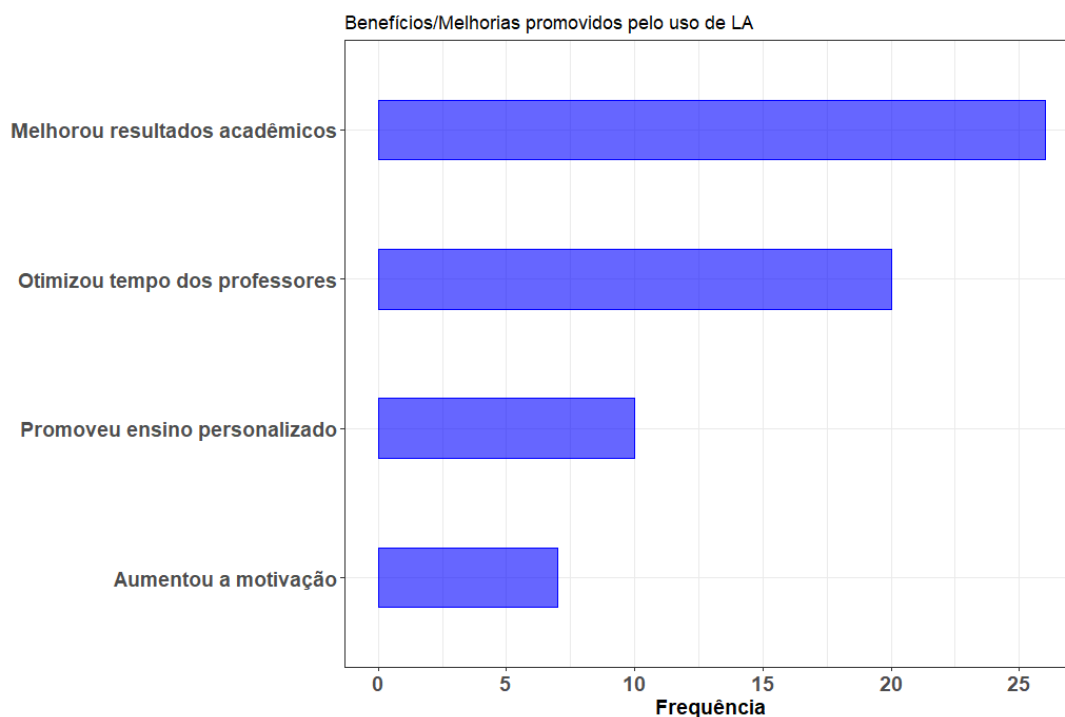


Figura 4: Benefícios/melhorias promovidos pelo uso de Learning Analytics.

- RQ9: A quem se destinava, diretamente, o resultado da aplicação de LA?

Observando as tabelas (**Tabela 12**, **Tabela 13** e **Tabela 14**), a questão de pesquisa RQ9 é respondida, e podemos concluir que quando trata-se de nível de ensino, o resultado da aplicação de LA é bem abrangente (22 artigos / 59%), sendo aplicada frequentemente na graduação (14 artigos / 37%), pouco comum no ensino fundamental (1 artigo / 2%) e na pós-graduação (1 artigo / 2%) e em nenhum dos artigos foi aplicado para o ensino médio. A área do conhecimento foi geral/independente em 27 dos 38 artigos (71%), restrito para matemática em 4 artigos (11%), para

programação em 3 estudos (8%), para gramática em 2 artigos (5%) e tanto para engenharia como medicina foi aplicado em 1 estudo (2%), ou seja, Learning Analytics pode ser aplicada independentemente da área e/ou disciplina. Quanto ao direcionamento do estudo, em 28 dos 38 artigos (74%) foi para ambos, tanto professores como estudantes, em 7 artigos (18%) foi restrito à professores, e em 3 artigos (8%) o foco foram os alunos.

Tabela 12: Nível de ensino em que o estudo foi aplicado.

Nível de ensino	Quantidade	%
Abrangente	22	59
Graduação	14	37
Primário/fundamental	1	2
Pós-graduação/profissional	1	2
Médio	0	0

Tabela 13: Área do conhecimento.

Área do conhecimento	Quantidade	%
Geral/independente	27	72
Matemática	4	11
Programação	3	8
Gramática/linguagens	2	5
Engenharia	1	2
Medicina	1	2

Tabela 14: Direcionamento do estudo.

Direcionamento do estudo	Quantidade	%
Ambos	28	74
Professores	7	18
Estudantes	3	8



- RQ10: Qual a percepção dos usuários finais em relação ao uso de Learning Analytics?

A questão de pesquisa RQ10, pode ser respondida através da observação da **Tabela 15**, onde temos que as percepções dos usuários finais em relação aos artigos selecionados foram consideravelmente positivas, tendo muitas avaliações de uso útil/efetivo de Learning Analytics (20 artigos / 52%), percepções de uso satisfatório em 14 artigos (37%), neutra em 4 artigos (11%) e em nenhum dos artigos os usuários consideraram que o esforço para utilizar Learning Analytics gerou resultados insatisfatórios, ou seja, uma percepção trabalhosa.

Tabela 15: Percepções dos usuários finais.

Percepções	Quantidade	%
Útil/efetiva	20	52
Satisfatória	14	37
Neutra	4	11
Trabalhosa	0	0

#### 4. Discussão dos resultados obtidos

A partir dos resultados apresentados e da discussão acerca das formas de aplicação que é descrita a seguir, poderemos compreender o campo de Learning Analytics na atualidade e entender seu uso para cada proposta, principalmente, para ajudar a identificar os estudantes que estão enfrentando dificuldades em um determinado curso online (**RQ0**).

Inicialmente, dados os resultados da **RQ1**, temos os principais objetivos do uso de Learning Analytics, como sendo: a previsão de estudantes em risco, que foi a proposta de aplicação mais frequente, seguida da visualização de informações e a personalização e adaptação. Definido o propósito que será abordado, necessitamos de um ambiente de aprendizagem online, que pode ser um ambiente já popular, ou então um novo sistema pode ser desenvolvido para atender essa demanda (**RQ3**).

A partir dos resultados da **RQ4**, sabemos que o ambiente de aprendizagem utilizado precisa armazenar os dados das interações dos estudantes com os recursos presentes no sistema, os resultados das avaliações online e/ou as mensagens em discussões online. A coleta e pré-processamento (**RQ5**) destas informações, deve ser feita através dos logs do sistema, ou seja, dos dados históricos dos estudantes que ficam armazenados no banco de dados (Gašević et al., 2016). Caso seja utilizada a coleta em tempo real ao invés dos dados históricos, o sistema necessitará de algum plugin e/ou ferramenta capaz de capturar e analisar as interações dos estudantes no ambiente online de forma automática e em tempo real (Filvà et al., 2019).

Após a coleta e pré-processamento dos dados, dados os resultados da **RQ2**, estas informações relacionadas às interações com os recursos do sistema e os resultados das avaliações online devem ser analisadas através do uso de técnicas e algoritmos que tenham a capacidade de identificar tendências ou realizar análises preditivas, permitindo que os professores estejam preparados para tomar decisões pedagógicas com base nos dados dos estudantes (You, 2016), que podem ser históricos ou em tempo real.

Quando trata-se de dados históricos, é possível o uso de algoritmos/técnicas de menor complexidade, como: análise de correlação, análise de variância, análises descritivas e testes de hipótese (You, 2016), o que deve ser considerado/ponderado na escolha da estratégia de análise dos dados. Quando estamos tratando de dados em tempo real, as chances são de que o problema necessite de algoritmos/técnicas inteligentes (Inteligência Artificial), como: aprendizado de máquina, redes neurais e/ou análise de redes sociais (Akçapınar et al., 2019).

Os algoritmos/técnicas mencionados (**Tabela 8**), podem ser utilizados para diversos objetivos, como a previsão de estudantes em risco no curso/disciplina, visualização de informações a respeito da situação/status de aprendizagem do estudante, personalização e adaptação da aprendizagem (**Tabela 7**). Um destes algoritmos/técnicas, a análise de redes sociais, possui uma diferença em relação aos tipos de dados utilizados, pois seu foco é a coleta dos dados relacionados a mensagens em fóruns e ambientes de discussões online (Hernández-García et al., 2015), enquanto as técnicas estatísticas, o aprendizado de máquina, entre outros, utilizam os dados das interações com os recursos do ambiente de aprendizagem e os resultados das avaliações online.

A última etapa, é a tarefa de enviar os resultados da análise dos dados dos estudantes para os professores e/ou instrutores, conforme os resultados da **RQ6**. Podemos realizar esta tarefa com o apoio de um dashboard (Leeuwen et al., 2019) com visualizações do resumo das interações dos alunos com os recursos do ambiente e os principais indicadores de performance (KPIs - Key Performance Indicators) da aprendizagem (Bhatti et al., 2014). A partir do uso de dashboards, ou seja, da interação entre o educador e o sistema (**RQ7**), o trabalho dos professores na tomada de decisão, e no entendimento da situação atual dos seus alunos, pode ser facilitado, permitindo auxiliá-los a realizar adaptações do método de ensino durante o curso (Dermeval et al., 2018), fornecendo feedback personalizado, e em tempo real, através do diálogo e interação aluno-professor com mensagens e alertas no ambiente online. Tais ações visam evitar que os estudantes com rendimento baixo evadam o curso/disciplina (**RQ8**).

Também podemos realizar uma classificação dos algoritmos/técnicas mais utilizados para cada questão/objetivo no uso de Learning Analytics, tomando como base os resultados da sumarização dos dados dos artigos selecionados. Abaixo, podemos observar o ranking:

- Previsão de estudantes em risco de reprovação/desistência/abandono:
  - Técnicas:
    - 1 - *Análise estatística descritiva;*
    - 2 - *Inferência estatística;*
    - 3 - *Técnicas de aprendizado de máquina;*
  - Algoritmos:
    - 1 - *Análises de regressão;*
    - 2 - *Análises de variância;*
    - 3 - *Clusterização de dados;*
    - 4 - *Análises de correlação;*
    - 5 - *Random forest;*
    - 6 - *Naive bayes;*
    - 7 - *Algoritmos de redes neurais;*
    - 8 - *Algoritmos de árvores de decisão;*

### 9 - Testes de hipótese e teste de qui-quadrado.

- Visualização de informações sobre fóruns/chats em discussões online:
  - Técnicas:
    - 1 - *Análise de redes sociais;*
    - 2 - *Análise do discurso;*
    - 3 - *Análise de sentimentos;*
  - Algoritmos:
    - 1 - *Máquina de vetores de suporte;*
    - 2 - *Redes neurais recorrentes;*
    - 3 - *Random forest;*
    - 4 - *Naive bayes;*
- Personalização e adaptação:
  - Técnicas:
    - 1 - *Análise estatística descritiva;*
    - 2 - *Inferência estatística;*
    - 3 - *Multimodal learning analytics.*
  - Algoritmos:
    - 1 - *Análises de regressão;*
    - 2 - *Análises de variância;*
    - 3 - *Análises de correlação;*
    - 4 - *Testes de hipótese;*

É importante frisar que existem outros algoritmos, técnicas e ferramentas que não foram mencionados na classificação acima que podem ser utilizados para as diferentes propostas de Learning Analytics existentes.

Vale ressaltar, que nos últimos anos houve um grande aumento do interesse no problema do abandono/desistência, que afeta escolas e universidades de todo o mundo, principalmente quando trata-se da educação online, adotada em instituições de todo o mundo devido à pandemia da Covid-19. Segundo o último Censo Ead Brasil (2019), instituições de ensino que ofereciam cursos a distância chegavam a ter uma taxa de evasão de até 50%, sendo ainda maior quando trata-se de MOOCs, seja por falta de apoio, dificuldades com a plataforma e/ou outros motivos (Aldowah et al., 2020). Uma das melhores propostas de solução deste problema da evasão em cursos online é a previsão dos estudantes em risco de reprovação e abandono/desistência (Hlosta et al., 2017), que foi listado como o principal objetivo de Learning Analytics (contando com um número maior de técnicas e algoritmos na classificação acima. Vide Seção 3 para mais informações).

Existem diversos estudos que pesquisam a área de Learning Analytics, incluindo outros mapeamentos e revisões sistemática da literatura, porém, nosso artigo possui contribuições que não estão presentes em outros trabalhos. Comparando o presente estudo com as produções científicas de Shane Dawson (2014) e de Barbara Moissa (2015), que são muito relevantes para o

campo de Learning Analytics, podemos identificar diferenças na abordagem, nas questões de pesquisa e nos objetivos dos artigos.

O estudo de Shane Dawson (2014) realiza uma avaliação geral do campo de Learning Analytics, analisando tanto estudos conceituais como os que incluem pesquisa empírica, mas nos seus resultados a grande maioria era de estudos teóricos. Diferente do presente estudo, que selecionou e avaliou estudos que abordam a aplicação prática de Learning Analytics em ambientes online de aprendizagem. A principal contribuição do artigo de Shane Dawson foi a análise do campo de LA adotando um esquema dividido em categorias: estudos com método qualitativo, com método quantitativo, métodos mistos e estudos com discussão de um tópico ou campo específico de forma mais geral.

O mapeamento sistemático de Barbara Moissa (2015) avalia estudos que abordam a aplicação prática de Learning Analytics, porém, com foco na análise dos resultados no contexto de Massive Open Online Courses (MOOCs). A partir dos resultados do estudo, o autor identificou tipos de dados utilizados em Learning Analytics como um todo que não se aplicam em plataformas MOOC, além de concluir que os artigos que envolvem o uso desse tipo de ambiente não costumam abordar previsões e intervenções, enquanto o foco do nosso mapeamento foi justamente na identificação de estudos que focam nessa abordagem, para responder a principal questão de pesquisa.

A principal contribuição do presente estudo foi a análise e extração de dados dos artigos da área para entender como podemos auxiliar professores na identificação de estudantes em risco, promovendo melhoria dos resultados acadêmicos em ambientes online de aprendizagem, ou seja, possui objetivos e abordagens diferentes em relação aos dois artigos mencionados e diversos outros relevantes para o campo de Learning Analytics.

Por fim, após a discussão apresentada anteriormente e a apresentação dos resultados do estudo, foi possível entender que Learning Analytics pode contribuir para a melhoria dos resultados acadêmicos dos estudantes, para a redução da evasão e para a otimização do tempo que os professores dedicam a sua atuação profissional em ambientes online de aprendizagem (**Figura 6**), além de outros benefícios que foram apresentados durante o mapeamento sistemático da literatura, independente da área do conhecimento do curso/disciplina (**Tabela 13**).

## 5. Conclusões e considerações finais

O objetivo principal do presente estudo, foi a obtenção de uma visão geral do campo de Learning Analytics para gerar orientações de uso em um ambiente de aprendizagem online de forma efetiva, para responder ao problema de pesquisa (**RQ0**: De que forma Learning Analytics é utilizada para ajudar a identificar os estudantes que estão enfrentando dificuldades em um curso online?) e às demais questões de pesquisa. Para tanto, foi conduzido um mapeamento sistemático da literatura para selecionar os principais artigos de Learning Analytics em diferentes aplicações, verificar o conteúdo destes estudos e, conseqüentemente, sumarizar e analisar os resultados para cumprir o objetivo da pesquisa.

O resultado do mapeamento sistemático nos permitiu responder nossas questões de pesquisa: a partir da **RQ1**, nossa conclusão foi que o principal objetivo do uso de Learning Analytics é para previsão de estudantes em risco; Para a **RQ2**, foi possível observar que os métodos mais utilizados foram as regressões e demais técnicas estatísticas, além dos algoritmos de aprendizado de máquina; Com base na **RQ3**, os sistemas mais utilizados foram o Moodle ou LMS/MOOC próprio; A partir da **RQ4** e **RQ5**, nossa conclusão é que os dados mais utilizados são os registros das interações do ambiente online, coletados/pré-processados através da extração dos logs; A

partir da **RQ6**, podemos concluir que as formas de visualização dos dados dos estudantes mais utilizadas são gráficos e tabelas, geralmente exibidos em um dashboard; Com base na **RQ7**, nossa conclusão foi que na maioria dos estudos houve a colaboração entre a inteligência humana e artificial; Por fim, a partir das questões **RQ8**, **RQ9** e **RQ10**, concluímos que a aplicação de Learning Analytics é efetiva em várias áreas do conhecimento, gerando diversos benefícios e melhorias na aprendizagem online.

Com base nos resultados encontrados a partir das questões de pesquisa, foi possível concluir que Learning Analytics pode ser utilizada para auxiliar na solução de diversos problemas da aprendizagem a distância, e principalmente para ajudar os professores a identificar e auxiliar os estudantes que estão enfrentando dificuldades em cursos/disciplinas online (**RQ0**), ou seja, realizar a previsão dos estudantes em risco de desistência/abandono. A contribuição do presente estudo, além de responder às questões de pesquisa, foi a atualização da área de LA, apresentando as diferentes formas de aplicação que têm sido utilizadas recentemente, uma classificação dos algoritmos/técnicas mais utilizados para cada problema abordado/proposto e os benefícios do uso em ambientes de aprendizagem online para professores e estudantes de diversas áreas do conhecimento e níveis de ensino, promovendo a melhoria do ensino a distância.

Apesar do presente estudo demonstrar os benefícios de utilizar Learning Analytics em ambientes de aprendizagem online, principalmente para auxiliar professores na identificação de estudantes em risco, algumas limitações devem ser observadas. Para que a aplicação de LA seja viável, faz-se necessário ter dados em quantidade e diversidade, ou seja, os alunos precisam interagir com os diversos recursos do ambiente online para que os algoritmos utilizados consigam prever seu respectivo desempenho no curso. Além disso, através dos estudos selecionados não foi possível constatar as diferenças entre os resultados da aplicação de algoritmos supervisionados e não-supervisionados nos dados coletados. Por fim, é necessário ter cautela na interpretação e precisão dos resultados dos algoritmos de análise, para que os Dashboards do ambiente apresentem informações confiáveis e úteis para os professores.

Em futuros trabalhos, os resultados encontrados no presente estudo poderão ser utilizados para implementar Learning Analytics em ambientes de aprendizagem online existentes ou até no desenvolvimento de um novo ambiente, cumprindo o objetivo de identificar antecipadamente estudantes em risco de reprovação e desistência, com visualizações valiosas que ajudem os professores em seus cursos online, aplicando as boas práticas que foram apresentadas a partir das respostas das perguntas de pesquisa, contribuindo assim, para o crescimento da área de Learning Analytics.

## Referências Bibliográficas

- Akçapınar, G., Hasnine, M. N., Majumdar, R., Flanagan, B., & Ogata, H. (2019). Developing an early-warning system for spotting at-risk students by using eBook interaction logs. *Smart Learning Environments*, 6(1), 4. doi: [10.1186/s40561-019-0083-4](https://doi.org/10.1186/s40561-019-0083-4) [GS Search]
- Aldowah, H., Al-Samarraie, H., Alzahrani, A. I., & Alalwan, N. (2020). Factors affecting student dropout in MOOCs: A cause and effect decision-making model. *Journal of Computing in Higher Education*, 32(2), 429-454. doi: [10.1007/s12528-019-09241-y](https://doi.org/10.1007/s12528-019-09241-y) [GS Search]
- Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, 20(2), 13-29. doi: [10.24059/olj.v20i2.790](https://doi.org/10.24059/olj.v20i2.790) [GS Search]

- Bhatti, M. I., Awan, H. M., & Razaq, Z. (2014). The key performance indicators (KPIs) and their impact on overall organizational performance. *Quality & Quantity*, 48(6), 3127-3143. doi: [10.1007/s11135-013-9945-y](https://doi.org/10.1007/s11135-013-9945-y) [GS Search]
- BR, Censo Ead. (2019). Relatório analítico da aprendizagem a distância no Brasil. url: [Censo EAD](https://censoead.gov.br/)
- Clow, D. (2012, April). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 134-138). doi: [10.1145/2330601.2330636](https://doi.org/10.1145/2330601.2330636) [GS Search]
- Dawson, S., Gašević, D., Siemens, G., & Joksimovic, S. (2014, March). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 231-240). doi: [10.1145/2567574.2567585](https://doi.org/10.1145/2567574.2567585) [GS Search]
- Dermeval, D., Paiva, R., Bittencourt, I. I., Vassileva, J., & Borges, D. (2018). Authoring tools for designing intelligent tutoring systems: a systematic review of the literature. *International Journal of Artificial Intelligence in Education*, 28(3), 336-384. doi: [10.1007/s40593-017-0157-9](https://doi.org/10.1007/s40593-017-0157-9) [GS Search]
- Filvà, D. A., Forment, M. A., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2019). Clickstream for learning analytics to assess students' behavior with Scratch. *Future Generation Computer Systems*, 93, 673-686. doi: [10.1016/j.future.2018.10.057](https://doi.org/10.1016/j.future.2018.10.057) [GS Search]
- Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, 28, 68-84. doi: [10.1016/j.iheduc.2015.10.002](https://doi.org/10.1016/j.iheduc.2015.10.002) [GS Search]
- Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64-71. doi: [10.1007/s11528-014-0822-x](https://doi.org/10.1007/s11528-014-0822-x) [GS Search]
- Hernández-García, Á., González-González, I., Jiménez-Zarco, A. I., & Chaparro-Peláez, J. (2015). Applying social learning analytics to message boards in online distance learning: A case study. *Computers in Human Behavior*, 47, 68-80. doi: [10.1016/j.chb.2014.10.038](https://doi.org/10.1016/j.chb.2014.10.038) [GS Search]
- Hlosta, M., Zdrahal, Z., & Zendulka, J. (2017, March). Ouroboros: early identification of at-risk students without models based on legacy data. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 6-15). doi: [10.1145/3027385.3027449](https://doi.org/10.1145/3027385.3027449) [GS Search]
- Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Keele University and Durham University Joint Report. url: [EBSE 2007-001](https://www.keele.ac.uk/department_of_computer_science/research/research_reports/2007-001/) [GS Search]
- Moissa, Barbara & Gasparini, Isabela & Kemczinski, Avaniilde. (2015). A Systematic Mapping on the Learning Analytics Field and Its Analysis in the Massive Open Online Courses Context. *International Journal of Distance Education Technologies*. 13. 1-24. doi: [10.4018/IJDET.2015070101](https://doi.org/10.4018/IJDET.2015070101) [GS Search]
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008, June). Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12* (pp. 1-10). doi: [10.14236/ewic/EASE2008.8](https://doi.org/10.14236/ewic/EASE2008.8) [GS Search]

- Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society*, 17(4), 117-132. url: [jeductechsoci.17.4.117](http://jeductechsoci.17.4.117) [GS Search]
- Siemens, G., & Baker, R. S. D. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). doi: [10.1145/2330601.2330661](https://doi.org/10.1145/2330601.2330661) [GS Search]
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380-1400. doi: [10.1177/0002764213498851](https://doi.org/10.1177/0002764213498851) [GS Search]
- Tseng, S. F., Tsao, Y. W., Yu, L. C., Chan, C. L., & Lai, K. R. (2016). Who will pass? Analyzing learner behaviors in MOOCs. *Research and Practice in Technology Enhanced Learning*, 11(1), 1-11. doi: [10.1186/s41039-016-0033-5](https://doi.org/10.1186/s41039-016-0033-5) [GS Search]
- van Leeuwen, A., Rummel, N., & Van Gog, T. (2019). What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations?. *International Journal of Computer-Supported Collaborative Learning*, 14(3), 261-289. doi: [10.1007/s11412-019-09299-x](https://doi.org/10.1007/s11412-019-09299-x) [GS Search]
- Xing, W., Chen, X., Stein, J., & Marcinkowski, M. (2016). Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization. *Computers in human behavior*, 58, 119-129. doi: [10.1016/j.chb.2015.12.007](https://doi.org/10.1016/j.chb.2015.12.007) [GS Search]
- You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education*, 29, 23-30. doi: [10.1016/j.iheduc.2015.11.003](https://doi.org/10.1016/j.iheduc.2015.11.003) [GS Search]
- Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, 27, 44-53. doi: [10.1016/j.iheduc.2015.05.002](https://doi.org/10.1016/j.iheduc.2015.05.002) [GS Search]