



Para subsidiar a extração de dados, um formulário de apoio foi definido no Google Docs. A partir dessa coleta de dados, os dados foram organizados em tabelas e gráficos no sentido de caracterizar os resultados obtidos. A Seção 4 apresentará a síntese desses resultados.

3.6. Execução do mapeamento sistemático

Considerando o processo de condução, por meio das bases de dados foi possível identificar 101 estudos na busca inicial. Desse total, 15 estudos foram incluídos ao se fazer a leitura de título, resumo e palavras-chave. Os 15 estudos restantes foram lidos na íntegra e 10 deles foram selecionados por extração de dados. A Tabela 3 contém uma síntese da execução. Vale salientar que antes da extração de dados, os estudos que foram referenciados pelos 10 estudos primários selecionados também foram analisados, usando *snowballing* e, a partir disso foram incluídos 12 novos estudos. Portanto, 22 estudos constituíram a amostragem de trabalhos que discute sobre a temática abordada neste trabalho.

Tabela 3. Resumo da condução do mapeamento

	Fase de seleção		Fase de extração	
	Incluídos	Excluídos	Incluídos	Excluídos
ACM Digital Library	0	1	0	0
Engineering Village	0	12	0	0
IEEE Xplore Digital Library	0	0	0	0
ScienceDirect	1	3	0	1
Scopus	10	62	8	2
Web of Science	4	8	2	2

4. Resultados

Esta seção foi estruturada de modo a responder cada QP deste mapeamento sistemático. As respostas foram organizadas a partir de dados coletados nos estudos que são listados na Tabela 4¹⁰. Nesta tabela, foi incluído um identificador (ID) para cada estudo primário, que será utilizado para apontar a origem de cada resultado apresentado.

Quais são os indicadores de evasão no contexto da EAD?

Para responder esta questão de pesquisa, os estudos foram analisados individualmente e classificados de acordo com o indicador de evasão que se encontrou. Todos os indicadores de evasão identificados nos estudos primários estão listados na Tabela 5. A primeira coluna da tabela relaciona os indicadores de evasão, a segunda coluna lista os estudos primários que tratam cada um deles. Destacam-se os indicadores comportamentais¹¹ que foram utilizadas em 17 estudos primários, em seguida o indicador de desempenho¹² empregado em 15 estudos primários.

Qual a relação dos indicadores de evasão com os recursos do AVA?

Cada indicador de evasão identificado na QP1 foi analisado e então relacionado aos recursos do AVA, como pode ser observado na Tabela 6, primeira e segunda colunas. Posteriormente, a terceira coluna lista os estudos primários correspondentes a cada categoria. Salienta-se que a maioria dos trabalhos analisados abordam o indicador comportamental. Observou-se

¹⁰As referências dos estudos estão disponíveis em: <<https://bit.ly/322p6oT>>.

¹¹Os indicadores comportamentais são relacionados a fatores relativos à aprendizagem ativa, motivação e envolvimento dos alunos, como acesso ao curso, aos materiais lidos e sua evolução com a programação do curso (MAZZA; DIMITROVA, 2007).

¹²Os indicadores de desempenho estão relacionados ao desempenho geral dos alunos, com base em seu desempenho nas atividades (questionário, tarefas, fóruns de discussão, dentre outros) (MAZZA; DIMITROVA, 2007).

¹³Os indicadores demográficos são relacionados a fatores relativos a dados históricos como sexo, idade, estado civil, ano de graduação, status anterior de graduação, localização, profissional (HLIOUI; ALOUI; GARGOURI, 2020).



Tabela 4. Lista completa dos estudos primários

Título do estudo	ID do estudo
A report-type plugin to indicate dropout risk in the Virtual Learning Environment Moodle	[E1]
An Infographics-based Tool for Monitoring Dropout Risk on Distance Learning in Higher Education	[E2]
Classification and predictive analysis of educational data to improve the quality of distance learning courses	[E3]
Co-embeddings for Student Modeling in Virtual Learning Environments	[E4]
CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses	[E5]
Efficient computation of key performance indicators in a distance learning university	[E6]
Identificação de Perfis de Evasão e Mau Desempenho para Geração de Alertas num Contexto de Educação a Distância	[E7]
Information management processes for extraction of student dropout indicators in courses in distance mode	[E8]
Learning Analytics in Practice: Providing E-Learning Researchers and Practitioners with Activity Data	[E9]
Learning to Identify At-Risk Students in Distance Education Using Interaction Counts	[E10]
Multi-agent System Based on Fuzzy Logic for Elearning Collaborative System	[E11]
OULAD MOOC Dropout and Result Prediction using Ensemble, Deep Learning and Regression Techniques	[E12]
Pedagogical monitoring as a tool to reduce dropout in distance learning in family health	[E13]
Predicting Students Success in Blended Learning—Evaluating Different Interactions Inside Learning Management Systems	[E14]
Predicting students' final performance from participation in on-line discussion forums	[E15]
Students' Success Predictive Models Based on Selected Input Parameters Set	[E16]
Um estudo do uso de contagem de interações semanais para predição precoce de evasão em educação a distância	[E17]
Um modelo preditivo para diagnóstico de evasão baseado nas interações de alunos em fóruns de discussão	[E18]
Uma Abordagem Genérica de Identificação Precoce de Estudantes com Risco de Evasão em um AVA utilizando Técnicas de Mineração de Dados	[E19]
Understanding Learner Engagement in a Virtual Learning Environment	[E20]
Using data mining as a strategy for assessing asynchronous discussion forums	[E21]
What can completion time of quizzes tell us about students' motivations and learning strategies?	[E22]

Tabela 5. Indicadores de evasão no contexto da EAD

Indicador	ID do estudo
Comportamental	[E1], [E2], [E3], [E5], [E7], [E8], [E9], [E10], [E11], [E12], [E14], [E15], [E16], [E17], [E18], [E20], [E21]
Demográfico ¹³	[E6], [E7], [E13], [E16], [E20]
Desempenho	[E1], [E2], [E4], [E5], [E9], [E12], [E14], [E15], [E16], [E17], [E18], [E19], [E20], [E21], [E22]

que o comportamento do aluno pode ser associado a interações em chat, conteúdo, fórum, mensagem, questionário, tarefa e videoconferência em AVAS e que o desempenho do aluno poderá ser medido pela participação em fórum, envio de questionário e envio de tarefas já relacionados ao indicador comportamental. Em contrapartida, o indicador demográfico tem relação apenas com o perfil do estudante. O estudo primário [E8] não identifica o recurso do AVA de onde provêm os dados, portanto este estudo não foi inserido na tabela.

Qual a relação dos indicadores de evasão com os dados que persistem no AVA?

Para responder sobre a relação dos indicadores de evasão com os dados que persistem no AVA, um esquema de classificação foi organizado, neste caso por indicador–dado. Pode-se observar que no indicador demográfico foram utilizados somente dados históricos (*e.g.*, sexo, idade, estado civil, ano de graduação, status anterior de graduação, localização, profissional) que são encontrados no recurso perfil do AVA. O indicador comportamental



Tabela 6. Relação dos indicadores de evasão com os recursos do AVA

Indicador	Recurso	ID do estudo
Demográfico	Perfil	[E13], [E20], [E16], [E6], [E7]
	Questionário	[E2], [E3], [E7], [E10], [E14], [E20]
	Fórum	[E1], [E2], [E3], [E5], [E7], [E9], [E10], [E11], [E12], [E14], [E15], [E17], [E18], [E20], [E21]
Comportamental	Mensagem	[E1], [E7], [E9], [E10], [E11], [E12], [E14], [E17]
	Tarefa	[E2], [E3], [E7], [E10], [E16]
	Conteúdo	[E2], [E10], [E12], [E17]
	Videoconferência	[E5]
	Chat	[E5]
Desempenho	Questionário	[E1], [E2], [E5], [E12], [E17], [E20], [E22]
	Fórum	[E9], [E15], [E18], [E21]
	Tarefa	[E2], [E4], [E5], [E12], [E14], [E16], [E17], [E19], [E20]

relaciona-se às interações realizadas em recursos e atividades do AVA, tais como *logs*, recursos acessados e atividades acessadas. O indicador de desempenho é obtido por meio de notas em atividades, participação em fórum e questionário do AVA. Constatou-se que os dados provindos de questionários, fóruns e atividades podem ser utilizados como medida dos indicadores comportamental e de desempenho, conforme a Tabela 7.

Tabela 7. Relação dos indicadores com os dados do AVA

Indicador	Dado	ID do estudo
Comportamental	Interações	[E1], [E2], [E3], [E5], [E7], [E8], [E9], [E10], [E11], [E12], [E14], [E15], [E16], [E17], [E18], [E20], [E21]
Demográfico	Dados históricos	[E6], [E7], [E13], [E16], [E20]
Desempenho	Nota	[E1], [E2], [E4], [E5], [E9], [E12], [E14], [E15], [E16], [E17], [E18], [E19], [E20], [E21], [E22]

Qual a relação dos indicadores de evasão com o público-alvo?

A relação dos indicadores de evasão com o público alvo é ilustrada na Tabela 8. Para identificar o público-alvo, foi considerado o nível de escolaridade. Constatou-se que a maioria dos estudos foram aplicados no contexto de cursos de graduação. Observou-se que o indicador demográfico só é considerado em cursos de graduação, já os indicadores comportamental e de desempenho são considerados em cursos de graduação e técnico (médio). As pesquisas de [E1], [E3] e [E20] não identificaram o perfil do público alvo, assim, não foram incluídas na Tabela 8.

Tabela 8. Relação dos indicadores de evasão com o público-alvo

Indicador	Nível educacional	ID do estudo
Comportamental	Graduação	[E11], [E14], [E15], [E16]
Comportamental e desempenho		[E2], [E4], [E5], [E7], [E9], [E10], [E11], [E12], [E16], [E21]
Demográfico		[E6], [E13]
Desempenho		[E19], [E22]
Comportamental	Técnico	[E8]
Comportamental e desempenho		[E17], [E18]

Os indicadores são usados em algum modelo de predição de evasão?

Conforme mencionado no decorrer deste trabalho, modelos têm sido construídos para prever alunos com risco de evasão. Considerando a temática deste estudo, foi investigado se no contexto dos estudos primários selecionados, modelos de predição de evasão foram estabelecidos. Constatou-se que 16 dos estudos primários ([E2], [E3], [E4], [E7], [E9], [E10],



[E11], [E12], [E14], [E15], [E16], [E17], [E18], [E19], [E20], [E21]) utilizaram os indicadores de evasão em modelos de predição. Seis (6) estudos primários ([E1], [E5], [E6], [E8], [E13], [E22]) abordaram os indicadores, mas não mencionaram sobre o uso dos mesmos em modelo de predição de evasão.

5. Conclusões

Neste trabalho foi apresentado o planejamento, a condução e os resultados de um mapeamento sistemático sobre o problema de evasão no contexto da EaD, em que se buscou elencar indicadores de evasão no contexto da EAD e sua relação com recursos e dados de AVAs. A partir dos resultados, constatou-se que os indicadores de evasão abordados nas pesquisas foram o indicador comportamental, indicador de desempenho e o indicador demográfico. Os dados coletados pelos AVAs, utilizados pelos indicadores, provêm de recursos como participação em fórum, envio de tarefas, interações em chat, respostas em questionários, dentre outros. Constatou-se ainda que a grande maioria dos estudos utilizaram os indicadores citados em modelos de predição de evasão e que a maioria dos estudos foram aplicados no contexto da graduação.

Espera-se que este estudo de mapeamento não sirva apenas para destacar os principais indicadores de evasão e sua relação com dados e recursos de AVAs, mas sirva para atrair pesquisadores e profissionais para descobrir um corpo de conhecimento que identifica principais indicadores de evasão para compor modelos de predição a fim de mitigar a evasão na EaD.

Além disso, como sugestão de trabalho futuro pode-se: verificar qual (quais) indicadores podem ter maior impacto na predição; identificar se há necessidade de utilizar todos os indicadores em modelos de predição; o porquê de um indicador utilizar dados que provêm de mais de um recurso do AVA; o porquê de um indicador utilizar dados de diferentes tipos (estruturado e não estruturado); investigar quais algoritmos são usados e como são construídos os modelos de predição relacionado os indicadores estudados; identificar se há um tratamento, considerando o dado usado com o algoritmo selecionado no modelo de predição, dentre outros.

Agradecimentos

Os autores gostariam de agradecer a CAPES - Código de Financiamento 001.

Referências

ABED. *Censo EAD.BR: relatório analítico da aprendizagem a distância no Brasil 2018 [livro eletrônico]/[organização] ABED – Associação Brasileira de Educação a Distância*. Curitiba: Intersaberes, 2019.

AGUDO-PEREGRINA Ángel F. *et al.* Can we predict success from log data in vles? classification of interactions for learning analytics and their relation with performance in vle-supported f2f and online learning. *Computers in Human Behavior*, v. 31, p. 542 – 550, 2014.

BALAKRISHNAN, G. *Predicting Student Retention in Massive Open Online Courses using Hidden Markov Models*, 2013. Available at: <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-109.html>.

BRITO, D. M. de *et al.* Identificação de estudantes do primeiro semestre com risco de evasão através de técnicas de data mining. In: *Congresso Internacional de Informática Educativa*, 2015. p. 459–463.



- CHEN, Y.; ZHANG, M. Mooc student dropout: Pattern and prevention. In: *ACM Turing 50th Celebration Conference*, 2017. p. 1–6.
- DIGIAMPIETRI, L. A.; NAKANO, F.; LAURETTO, M. d. S. Mineração de dados para identificação de alunos com alto risco de evasão: Um estudo de caso. *Revista de Graduação USP*, v. 1, n. 1, p. 17–23, 2016.
- HLIOUI, F.; ALOUI, N.; GARGOURI, F. Understanding learner engagement in a virtual learning environment. In: ABRAHAM, A. *et al.* (Ed.). *Intelligent Systems Design and Applications*. Cham: Springer International Publishing, 2020. p. 709–719.
- KHANNA, L.; SINGH, S. N.; ALAM, M. Educational data mining and its role in determining factors affecting students academic performance: A systematic review. In: *India International Conference on Information Processing*, 2016. p. 1–7.
- KIKA, C. A. *Supporting student experience management with learning analytics in the UK higher education sector*. Tese (Doutorado) — University of Bedfordshire, Inglaterra, 2018.
- KIZILCEC, R. F.; PÉREZ-SANAGUSTÍN, M.; MALDONADO, J. J. Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education*, v. 104, p. 18 – 33, 2017.
- LIMA, M. N.; ROBERTA, A. d. A. Educational datamining: A study of the factors that cause school dropout in higher education institutions in brazil. *RENOTE*, v. 18, n. 1, 2020.
- MACFADYEN, L. P.; DAWSON, S. Mining lms data to develop an “early warning system” for educators: A proof of concept. *Computers & education*, v. 54, n. 2, p. 588–599, 2010.
- MARQUES, L. T. *et al.* Mineração de dados auxiliando na descoberta das causas da evasão escolar: Um mapeamento sistemático da literatura. *RENOTE*, v. 17, n. 3, p. 194–203, 2019.
- MAZZA, R.; DIMITROVA, V. Coursevis: A graphical student monitoring tool for supporting instructors in web-based distance courses. *International Journal of Human-Computer Studies*, v. 65, n. 2, p. 125 – 139, 2007.
- MORAIS, A. M. de. *Abordagem avaliativa multidimensional para previsão da evasão do discente em cursos on-line*. Tese (Doutorado) — Universidade Federal de Campina Grande, Paraíba, Brasil, 2018.
- PETERSEN, K. *et al.* Systematic mapping studies in software engineering. In: *International Conference on Evaluation and Assessment in Software Engineering*, 2008. p. 68–77.
- QUEIROGA, E.; CECHINEL, C.; ARAÚJO, R. Um estudo do uso de contagem de interações semanais para predição precoce de evasão em educação a distância. *Workshop do Congresso Brasileiro de Informática na Educação*, p. 1074–1083, 2015.
- ROMERO, C.; VENTURA, S.; GARCÍA, E. Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, v. 51, n. 1, p. 368–384, 2008.
- TAMADA, M. M.; NETTO, J. F. de M.; LIMA, D. P. R. de. Predicting and reducing dropout in virtual learning using machine learning techniques: A systematic review. In: *IEEE Frontiers in Education Conference*, 2019. p. 1–9.