

A utilização de Metodologias Ativas com suporte de Mineração de Dados Educacionais e Learning Analytics para a mitigação da evasão em EaD: um mapeamento sistemático da literatura

Title: Application of Active Methodologies to students prone to dropout identified through Educational Data Mining and Learning Analytics: a systematic mapping of the literature

Título: Aplicación de Metodologías Activas a estudiantes propensos a la deserción identificados a través de Minería de Datos Educativos y Analítica de Aprendizaje: un mapeo sistemático de la literatura

Tiago Luís de Andrade
Universidade do Estado de Mato Grosso (UNEMAT)
ORCID: [0000-0003-0934-4814](https://orcid.org/0000-0003-0934-4814)
tiago@unemat.br

Caroline Medeiros Martins de Almeida
Universidade do Vale do Rio dos Sinos (UNISINOS)
ORCID: [0000-0002-0445-5921](https://orcid.org/0000-0002-0445-5921)
carolinemalmeida@unisinors.br

Jorge Luis Victória Barbosa
Universidade do Vale do Rio dos Sinos (UNISINOS)
ORCID: [0000-0002-0358-2056](https://orcid.org/0000-0002-0358-2056)
jbarbosa@unisinors.br

Sandro José Rigo
Universidade do Vale do Rio dos Sinos (UNISINOS)
ORCID: [0000-0001-8140-5621](https://orcid.org/0000-0001-8140-5621)
rigo@unisinors.br

Resumo

Apesar da ampla adoção da Educação a Distância (EaD), os altos índices de evasão são motivos de preocupação de professores e gestores institucionais. Existem iniciativas para mitigação dessa situação, como a aplicação de técnicas de Mineração de Dados Educacionais (MDE) e Learning Analytics (LA) para a identificação de alunos propensos a essa situação. No entanto, embora efetivos nessa tarefa, carecem de mecanismos para a motivação dos estudantes e intervenção pedagógica dos professores, já que não apresentam propostas metodológicas para incentivar a aprendizagem dos identificados com risco de evadir, mitigando essa possibilidade. A utilização de Metodologias Ativas após a identificação dos alunos através das técnicas de MDE e LA pode constituir um mecanismo efetivo de prevenção da evasão na EaD, ampliando o potencial de engajamento e colaboração entre os alunos. Este artigo apresenta um Mapeamento Sistemático da Literatura com o objetivo de identificar as técnicas mais utilizadas de MDE e LA no contexto de evasão. Além disso, identificar a aplicação de Metodologias Ativas para mitigar a possibilidade de evadir nos cursos ofertados na Educação a Distância. Avaliamos 1103 artigos publicados no período de janeiro de 2015 a março de 2023. Os resultados indicam uma crescente aplicação de MDE e LA para identificar e mitigar a evasão de alunos na EaD. Entretanto, estudos com a utilização da estratégia pedagógica das Metodologias Ativas para minimizar esse problema e potencializar a permanência dos alunos são escassos.

Palavras-Chave: Metodologias Ativas; Mineração de Dados Educacionais; Learning Analytics; Evasão; Educação a Distância.

Abstract

Despite the widespread adoption of Distance Education, the high dropout rates concern teachers and institutional managers. There are initiatives to mitigate this situation, such as applying Educational Data Mining (EDM) and Learning Analytics (LA) techniques to identify students prone to this situation. However, although effective in this task, they lack mechanisms for student motivation and teachers' pedagogical intervention since they do not present methodological proposals to encourage the learning of those identified as at risk of dropout, mitigating this

Cite as: Andrade, T. L. de, Almeida, C. M. M. de, Barbosa, J. L. V. & Rigo, S. J. (2023). A utilização de Metodologias Ativas com suporte de Mineração de Dados Educacionais e Learning Analytics para a mitigação da evasão em EaD: um mapeamento sistemático da literatura. *Revista Brasileira de Informática na Educação*, 31, 1057-1088. <https://doi.org/10.5753/rbie.2023.3594>

possibility. The use of Active Methodologies after identifying students through EDM and LA techniques can be an effective mechanism for preventing dropout in Distance Education, expanding the potential for student engagement and collaboration. This article presents a Systematic Mapping of the Literature to identify the most used EDM and LA techniques in the context of dropout. In addition, identify the application of Active Methodologies to mitigate the possibility of dropout of courses offered in Distance Education. We evaluated 1103 articles published from January 2015 to March 2023. The results indicate a growing application of EDM and LA to identify and mitigate student dropout in Distance Education. However, studies using the pedagogical strategy of Active Methodologies to minimize this problem and enhance student retention are scarce.

Keywords: Active Methodologies; Educational Data Mining; Learning Analytics; Dropout; Distance Education.

Resumen

A pesar de la adopción generalizada de la Educación a Distancia (EaD), las altas tasas de deserción son motivo de preocupación para los docentes y administradores institucionales. Existen iniciativas para mitigar esta situación, como la aplicación de técnicas de Minería de Datos Educativos (MDE) y Analítica de Aprendizaje (AA) para identificar a los estudiantes propensos a esta situación. Sin embargo, aunque efectivos en esta tarea, carecen de mecanismos para la motivación de los estudiantes y la intervención pedagógica de los docentes, ya que no presentan propuestas metodológicas para incentivar el aprendizaje de aquellos identificados en riesgo de deserción, mitigando esta posibilidad. El uso de Metodologías Activas después de identificar a los estudiantes a través de técnicas MDE y AA puede ser un mecanismo eficaz para prevenir la evasión en EaD, ampliando el potencial de participación y colaboración entre los estudiantes. Este artículo presenta un Mapeo Sistemático de la Literatura con el objetivo de identificar las técnicas de MDE y AA más utilizadas en el contexto de la evasión. Además, identificar la aplicación de Metodologías Activas para mitigar la posibilidad de deserción de los cursos ofrecidos en Educación a Distancia. Evaluamos 1103 artículos publicados desde enero de 2015 hasta marzo de 2023. Los resultados indican una creciente aplicación de MDE y AA para identificar y mitigar la deserción estudiantil en EaD. Sin embargo, los estudios que utilizan la estrategia pedagógica de Metodologías Activas para minimizar este problema y mejorar la retención de los estudiantes son escasos.

Palabras clave: Metodologías Activas; Minería de Datos Educativos; Análisis de Aprendizaje; Evasión; Educación a Distancia.

1 Introdução

Com o avanço das Tecnologias de Informação e Comunicação (TICs) na educação, novas formas de compartilhamento de saberes e desenvolvimento de conhecimento estão sendo utilizadas para o processo de ensino e aprendizagem e, nesse contexto, a Educação a Distância (EaD) vem assumindo um papel importante na formação de pessoas. A EaD é uma modalidade de ensino que permite que professores e alunos possam estar geograficamente distantes, interagindo de forma síncrona e assíncrona para a construção do conhecimento através de Ambientes Virtuais de Aprendizagem – AVA (Heidrich et al., 2018; Ramos et al., 2018).

Periodicamente, pesquisas são realizadas para coletar informações sobre os cursos ofertados em EaD. No Brasil, o Censo Anual apresentou uma taxa média de evasão entre 26% e 50% em 2015 (Ramos et al., 2017). Em 2016, a taxa média de evasão foi entre 11% e 35% (Ramos et al., 2018); e, em 2018, foram encontrados indicativos de até 50%. Já em 2019, de acordo com o Censo 2019/2020, os índices variaram entre 26% e 50%; e, conforme o CensoEAD.BR – 2020/2021, os índices variaram até 26%.

Nota-se, portanto, que o elevado índice de evasão dos estudantes preocupa os gestores e professores das instituições de ensino, que buscam alternativas para identificar as situações passíveis de desistências e motivar os alunos a permanecerem nos estudos (Manhães et al., 2011; Kostopoulos et al., 2018a; Kostopoulos et al., 2018b).

Existem trabalhos que utilizam as técnicas de Mineração de Dados Educacionais (MDE, do inglês *Educational Data Mining - EDM*) e *Learning Analytics (LA)* para identificar alunos propensos a evadirem do curso (Almeida Neto & Castro, 2015; Liang et al., 2016; Queiroga et al., 2017; Ramos et al., 2018; Ortigosa et al., 2019; Tomasevic et al., 2020; Esteban et al., 2021; Tamada et al., 2022; Tran et al., 2023), no entanto, restritas a tal ação, incapazes de reverter esse

cenário (Widyahastuti & Tjhin, 2018). Sendo assim, a tomada de decisão para mitigar esse problema é, normalmente, dependente do professor, que necessita utilizar de metodologias educacionais para resgatar e incentivar o aluno a permanecer no curso. Uma alternativa seria a inclusão de estratégias motivadoras e colaborativas das Metodologias Ativas, definidas como abordagens pedagógicas em que os alunos participam como protagonistas do processo de aprendizagem, estimulados a se relacionar com os colegas e com o trabalho em equipe para o desenvolvimento de atividades, do senso crítico e da capacidade de argumentação, colaborando para o crescimento intelectual e a melhoria no desempenho dos envolvidos (Mattar, 2017; Mota & Rosa, 2018; Almeida et al., 2019).

Diante do exposto, este mapeamento sistemático visa verificar, entre o período de 2015 a 2023, o uso de Mineração de Dados Educacionais (MDE) e *Learning Analytics* (LA) para a identificação de alunos propensos à evasão e analisar a integração das Metodologias Ativas a esse processo, a fim de auxiliar o professor na ensinar e colaborar na mitigação do abandono escolar. Como objetivos específicos, pretende-se: (i) identificar técnicas, algoritmos e aplicações de MDE e LA voltadas para AVA no processo de predição e/ou monitoramento de alunos; e (ii) identificar aspectos referentes ao uso de Metodologias Ativas em plataformas educacionais integradas ao processo de MDE e LA para mitigar a evasão e potencializar a permanência na educação a distância.

Este artigo está organizado em cinco seções, além desse contexto introdutório. A seção 2 apresenta a fundamentação teórica com os conceitos relevantes de educação a distância, evasão na educação a distância, MDE e LA na educação a distância e, por fim, Metodologias Ativas e educação a distância. A seção 3 descreve a metodologia da pesquisa e as questões que permeiam este estudo. A seção 4 relata os principais resultados obtidos para cada questão de pesquisa. Por fim, a seção 5 apresenta as considerações finais deste trabalho.

2 Fundamentação Teórica

Esta seção apresenta os conceitos acerca dos temas desse artigo, a saber: Educação a Distância, evasão, MDE, LA e Metodologias Ativas.

2.1 Educação a Distância

A EaD vem ao longo dos anos se expandindo como modalidade de ensino, sobretudo no que se refere ao acesso a educação superior, assumindo um papel importante na formação e capacitação de pessoas.

Embora haja uma ampla discussão sobre o que é Educação a Distância, Ramos et al. (2018), Heidrich et al. (2018) e Brito (2019b) definem como uma modalidade de ensino no qual alunos e professores desenvolvem o processo de ensino e aprendizagem, de forma síncrona e assíncrona, sem estarem no mesmo lugar e no mesmo horário, apoiados pelo uso massivo das Tecnologias da Informação e Comunicação (TICs). Brito (2019b) complementa ao apontar elementos que estabelecem as características para a EaD: a separação espacial e temporal entre professor e aluno, o intenso aporte das TICs, a participação ativa de tutores, a aprendizagem autônoma e flexível dos alunos e a comunicação ativa entre os atores professor, tutor e aluno.

De acordo com o último Censo da Educação Superior no Brasil, realizado em 2021, observa-se o aumento considerável dos números de vagas ofertadas nos cursos, ingressos e matrículas realizadas nas instituições de ensino no Brasil nos últimos anos, tanto na rede pública quanto na privada, consolidando essa modalidade de ensino e a reafirmando a tendência de expansão ano a ano, já que, desde 2020, os cursos de graduação à distância no Brasil receberam mais alunos novos do que os presenciais.

Entre os fatores pode-se citar, por exemplo, a autonomia dos estudantes em realizar os seus estudos em qualquer lugar e o acesso aos ambientes virtuais de aprendizagem a qualquer hora, graças aos recursos tecnológicos existentes acessíveis atualmente, como Ambientes Virtuais de Aprendizagem (AVA) ou MOOC (do inglês *Massive Open Online Course* – Curso Online Aberto e Massivo), plataformas educacionais *on-line* de fácil manuseio que proporcionam ao professor a inserção de materiais e vídeos para o aprendizado do aluno, bem como permitem a administração de cursos com a oferta de vários serviços interativos, como fóruns de discussão, *lives*, *chats*, dentre outros (Brito, 2019b).

Nota-se que o uso computacional e o alto nível de mediação digital na educação vêm gerando um grande volume de dados que podem servir como matéria prima para pesquisas (Cambruzzi et al., 2015). Dados como registros de acessos, interações e mensagens em fóruns, entre outros, podem ser analisados a fim de avaliar o comportamento dos alunos que apresentam sucesso e insucesso e auxiliar os instrutores a melhorar o desempenho de cada um (Waheed et al., 2020; Isidro et al., 2018; Ramos et al., 2018; Kostopoulos et al., 2019b).

Desse modo, os dados coletados com a consolidação da Educação a Distância e uso massivo dos ambientes virtuais nessa modalidade de ensino podem auxiliar professores e gestores institucionais para a inclusão e utilização de estratégias pedagógicas, com o objetivo de melhorar o ensino e a aprendizagem nas instituições educacionais.

2.2 Evasão na EaD

Atualmente, a Educação a Distância tem se consolidado como modalidade de ensino, entretanto a evasão de alunos é um dos principais problemas enfrentados pelas instituições educacionais públicas e privadas, motivo pelo qual tem gerado muitas pesquisas nacionais e internacionais sobre o tema (Manhães et al., 2011, Silva et al., 2015).

Em busca de uma definição sobre esse tema estudado por vários autores, a evasão pode ser definida como a interrupção do progresso acadêmico do estudante ao longo dos semestres, ou seja, a desistência definitiva do aluno em qualquer etapa do curso (Silva et al., 2015; Santos et al., 2021a).

No Brasil, a Associação Brasileira de Educação a Distância (ABED) elabora, anualmente, o CensoEAD.BR, um relatório analítico da aprendizagem a distância. De acordo com o último levantamento de evasão, realizado em 2020/2021, 85 Instituições de Ensino (IES) participaram da pesquisa e apontaram que a dificuldade financeira é o principal motivo que leva os alunos da graduação e pós-graduação a evadirem dos cursos ofertados nessa modalidade de ensino (ABED, 2022).

Ainda assim, em busca de identificar e compreender os fatores que contribuem para a desistência na EaD, Silva & Castro (2022) relacionam: (i) às instituições de ensino: quando se trata da falta de apoio administrativo e acadêmico, a infraestrutura do curso, a falta de interação e integração no AVA, a ausência de encontros presenciais e a falta de orientação e comunicação por parte de professores/tutores; e (ii) aos estudantes: como problemas pessoais e financeiros, falta de habilidades com recursos tecnológicos, falta de adaptação ao método de ensino, dificuldades de acesso à internet, falta de apoio familiar e da empresa em que trabalha e dificuldades para gerir o tempo.

Manhães et al. (2011) classificam a evasão sob três agrupamentos: (i) econômica, pois o aluno é impossibilitado de permanecer no curso por questões socioeconômicas; (ii) vocacional, uma vez que o aluno não se identifica com o curso; (iii) institucional, em que o abandono ocorre por fracasso nas disciplinas iniciais, deficiências prévias de conteúdos anteriores, inadequação aos métodos de estudo, dificuldades de relacionamento.

Diante disso, conhecer os motivos que levam o aluno a evadir é importante, no entanto, torna-se essencial buscar meios de identificar os alunos propensos a essa situação e, de alguma forma, mitigar essa possibilidade. Nesse caso, a utilização de métodos e ferramentas de análise de dados e aprendizagem tende a indicar fatores que podem levar o sucesso ou o fracasso dos alunos de cursos a distância (Silva et al., 2015).

Nesse cenário, técnicas de MDE e LA são apresentadas como uma alternativa para o tratamento e descoberta de conhecimento nas bases de dados, a partir das informações geradas pelos estudantes, com o uso das plataformas educacionais, tal qual o seu comportamento, interação e desempenho nas atividades avaliativas. Desse modo, MDE e LA vêm se estabelecendo como linha de pesquisa forte e consolidada para a predição de evasão dos alunos (Baker et al., 2011; Queiroga et al., 2019).

2.3 Mineração de Dados Educacionais e *Learning Analytics* em EaD

A Mineração de Dados Educacionais e *Learning Analytics* são áreas de pesquisa interdisciplinares que trabalham com o uso de Mineração de Dados e métodos estatísticos para explorar os dados educacionais dos alunos (Kostopoulos et al., 2019a).

Segundo Baker & Yacef (2009) e Kostopoulos et al. (2019a), MDE é caracterizada como uma área de pesquisa que busca desvendar conhecimento a partir de dados educacionais, formando um elemento integrante do processo de aprendizagem de estudantes, educadores e instituições de ensino. Similarmente, Silva et al. (2015) e Shafiq et al. (2022) afirmam que é uma abordagem indutiva interessante que cria modelos para descobrir informações em dados dos alunos que podem ser utilizadas na melhoria do processo de aprendizagem. Queiroga et al. (2022) atribuem à MDE o desenvolvimento e a adaptação de métodos que possam auxiliar na descoberta de informações em dados provenientes de múltiplas fontes e recursos educacionais. Assim, busca a compreensão de fatores que influenciam na aprendizagem, como o comportamento dos alunos e as situações que podem levá-los a aprender de forma mais eficaz.

LA, por sua vez, é um campo de pesquisa em rápido crescimento, focado principalmente na análise de dados para, a partir deles, contribuir com as intervenções dos professores e no processo de aprendizagem dos alunos, melhorando a qualidade do ensino e os resultados educacionais (Shafiq et al., 2022). Como afirmam Queiroga et al. (2022), LA faz uso de técnicas para o desenvolvimento de métodos que possam auxiliar na melhora da experiência no processo de aprendizagem e em todo o seu contexto. Brito et al. (2019a) afirmam que LA atraiu a atenção da comunidade científica que trabalha com tecnologias educacionais, pois fornece uma maneira mais eficaz para os professores acompanharem o desempenho dos alunos e o envolvimento nas plataformas educacionais.

Conforme destacado, as duas compartilham do mesmo objetivo: melhorar o processo de ensino e aprendizagem através do aperfeiçoamento dos processos de avaliação, da compreensão dos problemas da educação e do planejamento das intervenções (Siemens & d Baker, 2012).

No contexto da EaD, existem vários trabalhos publicados sobre MDE e LA com foco nessa modalidade de ensino. Entre os principais, Heidrich et al. (2018) fazem um diagnóstico da evasão baseando-se no comportamento do aluno da modalidade EaD. Rodrigues et al. (2018) trazem uma revisão dos últimos 20 anos sobre esse tema, com foco em estudos sobre interações entre os atores educacionais, monitoramento e avaliação do processo de ensino e aprendizagem, avaliação sobre as estratégias pedagógicas adotadas, riscos de evasão e recomendação de materiais educacionais. Queiroga et al. (2017) afirmam que a MDE e LA surgem como uma alternativa para o tratamento e a descoberta de conhecimento diante do grande volume de dados gerados e armazenados nos AVA.

Conforme exposto, entre as investigações possíveis de se realizar utilizando técnicas de MDE e LA em EaD, têm-se a verificação de potenciais riscos de evasão (Manhães et al., 2011; Queiroga et al., 2022), auxílio na formação de grupos (Correia & Pimentel, 2011) e análise do comportamento (Santos et al., 2015). Dessa forma, com a aplicação de técnicas de MDE e LA, pode-se realizar a previsão de desempenho acadêmico e estudo sobre os aspectos que indicam as possibilidades do aluno evadir. Essas descobertas são úteis para professores e gestores das instituições de ensino, já que podem auxiliá-los a tomar, de forma antecipada, decisões que aprimorem as práticas educacionais e instiguem a busca por estratégias pedagógicas que possam mitigar essas possibilidades e a ajudar os estudantes (Rodrigues et al., 2018).

Para aplicar as técnicas de MDE e LA, existem várias ferramentas capazes de auxiliar o professor, principalmente se utilizadas em ambientes virtuais para o processo de ensino e aprendizagem. No AVA Moodle¹ (*Modular Object-Oriented Dynamic Learning Environment*), ambiente de aprendizado virtual de código aberto mais amplamente utilizado para educação a distância em todo o mundo (Félix et al., 2018)), as ferramentas *Feedback*, *Survey* e *Questionnaire* fornecem informações relevantes sobre o processo.

Portanto, com a expansão da EaD e o uso massivo de ambientes como AVA e MOOC, investigar os dados gerados a partir da interação dos professores e alunos pode contribuir com as instituições de ensino no intuito de descobrir os problemas existentes na aprendizagem dos alunos e ajudar a melhorar a qualidade do ensino (Aldowah et al., 2019). Uma vez identificados, propor estratégias de resgate e motivação para que o aluno continue o seu percurso. Desse modo, é pertinente a utilização de estratégias pedagógicas que colaborem com o professor no processo de ensino e aprendizagem e ajude a mitigar a possibilidade de evasão.

2.4 Metodologias Ativas e EaD

O uso de Metodologias Ativas tem se expandido no ensino presencial e a distância, já que é possível encontrar diversas pesquisas no assunto que tratam da sua eficiência no processo de ensino e aprendizagem (Almeida et al., 2019; Mota & Rosa, 2018; Alves et al., 2020b).

De acordo com Acosta et al. (2018) e Li et al. (2018), as Metodologias Ativas englobam uma concepção do processo de ensino e aprendizagem que considera a participação efetiva dos alunos na construção da mesma, valorizando as diferentes formas pelas quais eles podem ser envolvidos para que aprendam melhor, em seu próprio ritmo, tempo e estilo. Para Mattar (2017), as Metodologias Ativas propõem uma educação que estimula a atividade (ao invés de passividade) por parte dos alunos. Em relação a atuação do professor, nota-se que nas Metodologias Ativas o seu papel também foi repensado, uma vez que deixa de ser mero transmissor do conhecimento para monitor, com o dever de criar ambientes de aprendizagem repletos de atividades diversificadas (Chandrasekaran et al., 2016; Mota & Rosa, 2018).

Como benefícios da utilização das Metodologias Ativas, Almeida et al. (2019) afirmam que elas têm potencial de tornar as aulas mais interessantes e modernas, indo ao encontro das necessidades recentes das universidades, podendo ser alternativas de estratégias educativas que facilitam a aprendizagem dos alunos. Nesse sentido, apropriar-se delas tem sido um desafio aos professores e estudantes, pois impõem a flexibilidade de abertura ao novo e a capacidade de aprender (Chandrasekaran et al., 2016).

Existem várias Metodologias Ativas, que se diferenciam à medida que definem suas estratégias, abordagens e técnicas, bem como o papel do professor e do aluno no processo de ensino e aprendizagem.

¹ <https://moodle.com/>

A Aprendizagem Baseada em Problemas – ABP ou *Problem Based Learning* – PBL usa-se o problema para ajudar os alunos a identificarem suas necessidades de aprendizagem, à medida que tentam entender, reunir, sintetizar e aplicar informações ao mesmo e começar a trabalhar efetivamente para aprender com os membros do grupo (Mattar, 2017; Fonseca & Mattar, 2018). A Aprendizagem Baseada em Projetos é uma metodologia na qual os alunos adquirem conhecimentos e habilidades trabalhando por um longo período de tempo na investigação, de forma a responder a uma questão, um problema ou um desafio autêntico, envolvente e complexo (Mattar, 2017).

O Ensino Híbrido ou *Blended Learning* mescla a aprendizagem via plataforma educacional e presencial em sala de aula, em que o aluno aprende por meio de aprendizagem *on-line*, sobre o qual tem algum controle em relação ao lugar, ao tempo, ao caminho e/ou ao ritmo, e parte em um local físico, supervisionado, longe de casa (Mattar, 2017). A Aprendizagem entre Pares ou *Peer Instruction* – PI visa construir antes, durante e após a aula, o conhecimento dos alunos a partir da interação, com menos aulas expositivas do professor e mais interativas com alunos (Araújo & Mazur, 2013).

A Sala de Aula Invertida ou *Flipped Classroom* aborda os conceitos teóricos a partir da responsabilização do aluno por estudar previamente os conceitos, a fim de desenvolver as atividades práticas na aula. Bergmann & Sams (2016) explicam que a proposta dessa metodologia é tornar aquilo que, tradicionalmente, é feito em sala de aula, ser executado em casa e, conseqüentemente, o que é feito como trabalho de casa, ser realizado em sala de aula. A Gamificação transforma situações de aulas em resolução de problemas, de forma criativa, por meio de elementos de jogos no processo de ensino (Schlemmer, 2014; Mattar, 2017).

Além das metodologias citadas, existem outras com potencial de direcionar os alunos a aprendizagens e que propiciam um desenvolvimento ativo, autônomo e com maior responsabilidade para o seu aprendizado, como *Just-in-Time-Teaching* ou Ensino sob Medida (Araújo & Mazur, 2013), Avaliação por Pares (Mattar, 2017; Hoffmann et al., 2020) e *Design Thinking* (Mattar, 2017).

O AVA *Moodle* possui ferramentas que possibilitam a utilização de Metodologias Ativas no processo de ensino e aprendizagem. *Moodle Block Game* e *Game* são exemplos que permitem a aplicação da metodologia Gamificação. *Wiki* e *Collabora* propiciam as atividades colaborativas entre alunos e professores. Basicamente, esses são exemplos de ferramentas, no entanto, existem outras que estão sendo aprimoradas ou desenvolvidas para atenderem a modalidade EaD.

3 Trabalhos Relacionados

Nesta seção são apresentados trabalhos relacionados a este que objetivam a investigação da aplicação de técnicas de MDE e LA para prever a possibilidade de evasão de alunos em cursos ofertados em EaD e também da utilização de Metodologias Ativas como estratégia de resgate e motivação dos estudantes nesta modalidade de ensino.

Em relação a recentes trabalhos, Esteban et al. (2021) realizaram um estudo de previsão do desempenho dos alunos baseado na execução de tarefas no ambiente virtual. Os resultados evidenciaram que, em um conjunto de 23 algoritmos de MDE utilizados, os atributos de interação do envio de tarefas foram importantes para a indicação e previsão dos alunos com riscos de evasão. Adnan et al. (2022) desenvolveram um modelo preditivo de Inteligência Artificial Explicável (*XAI - Explainable Artificial Intelligence*) que interpreta os dados do ambiente virtual e prevê o desempenho dos alunos em vários estágios da duração do curso. Já

Queiroga et al. (2022) apresentaram um estudo em dois contextos diferentes: presencial e a distância.

A respeito da EaD, o objetivo foi a predição de evasão precoce através da aplicação de técnicas de MDE em dados de interações dos estudantes, extraídos de quatro cursos técnicos oferecidos no AVA. Segundo os autores, diante de um contexto limitado de dados disponíveis, foi necessário adaptar metodologias para o processamento e modelagem, transformando as contagens de interações em uma linha do tempo de acompanhamento diário dos estudantes. Nesse processo, foi possível evidenciar o poder preditivo nas quatro primeiras semanas de curso.

Ainda sobre a aplicação de MDE e LA, Tamada et al. (2022) construíram um modelo de previsão de desempenho do aluno baseado em Aprendizado de Máquina usando dados acadêmicos e registros de um Sistema de Gestão de Aprendizagem (SGA), que correlaciona com o sucesso ou insucesso na conclusão do curso. Os resultados indicaram que aproximadamente 28% dos alunos que iniciaram bem não conseguiram concluir o curso, e apenas 2% dos que iniciaram com dificuldades nas avaliações conseguiram melhorar e concluir o curso. Especificamente sobre os algoritmos, *Random Forest* forneceu os melhores resultados, com 84,47% na pontuação dos experimentos. Por fim, Waheed et al. (2023) descreveram os resultados de um estudo que investigou o desempenho de uma técnica de aprendizado profundo chamada *Long Short-Term Memory (LSTM)* na previsão de alunos em risco de reprovação em um curso *on-line*. Os resultados apontaram uma precisão de 84,57% na identificação. Além disso, o estudo identificou os atributos mais importantes para a retenção dos alunos: o atraso na submissão de atividades e os erros nas respostas dos questionários avaliativos.

Nenhum dos trabalhos citados considerou integrar Metodologias Ativas após a previsão de evasão dos alunos. Nota-se a preocupação com os índices alarmantes de evasão no ensino a distância. Diante disso e da fundamentação teórica realizada, considera-se que conhecer as causas seja de extrema dificuldade, embora alguns fatores sejam citados e identificados por meio das técnicas de MDE e LA, como os que mais contribuem para a desistência: a interação social, os métodos de avaliação e a expectativa do aprendizado frustrado.

Especificamente sobre Metodologias Ativas, Chandrasekaran et al. (2016) realizou uma pesquisa para mensurar o conforto dos usuários ao utilizarem essa estratégia pedagógica no processo de aprendizagem. Os resultados demonstraram que 67% dos entrevistados sentiram-se confortáveis, uma vez que a metodologia ofereceu aos alunos a oportunidade de expressarem experiências individuais e compartilharem ideias em grupo, bem como atribuíram a si boa parte da responsabilidade em aprender. Nesse contexto, Lima & Siebra (2017) utilizaram Metodologias Ativas para mitigar a possibilidade de evasão e melhorar a permanência nos cursos oferecidos na educação a distância, no entanto, não fizeram o uso de MDE e LA para identificar os alunos propensos a evadir.

4 Metodologia

Esta seção apresenta um Mapeamento Sistemático da Literatura para o estudo dos trabalhos relacionados que, conforme Petersen et al. (2015), consiste nas seguintes etapas: a formulação de questões de pesquisa, a definição dos critérios de inclusão e exclusão dos estudos, a busca e seleção dos artigos a serem analisados, a avaliação dos estudos e, por último, a coleta dos dados. Em seguida, a etapa de análise e apresentação deu-se em formato de gráficos, números e descrições, fundamentando a interpretação dos resultados e discussões.

4.1 Questões de pesquisa

Para o processo de pesquisa foram definidas as questões apresentadas na tabela 1.

Tabela 1: Questões de pesquisa.

ID	Questão
Q1	Evasão em EaD tem sido objeto de estudo?
Q2	Quais técnicas de MDE e LA são utilizadas para prever, detectar, diagnosticar ou monitorar a evasão em EaD?
Q3	Quais ferramentas computacionais foram utilizadas para o estudo de evasão em EaD?
Q4	Quais algoritmos foram utilizados para o estudo de evasão em EaD?
Q5	Quais atributos foram utilizados no estudo de evasão em EaD?
Q6	Qual é o nível de escolaridade do público alvo nos estudos de evasão em EaD?
Q7	Metodologias Ativas foram utilizadas para mitigar a evasão em EaD?

4.2 Processo de pesquisa

Como estratégia para se descobrir os estudos relevantes, realizou-se uma busca automática em bases eletrônicas e uma busca manual, em conferência, para garantir que o maior número de estudos pudesse ser verificado. As seguintes bases de dados eletrônicas foram pesquisadas:

- *Institute of Electrical and Electronics Engineers (IEEE)*
- *Science Direct da Elsevier*
- *Association for Computing Machinery (ACM)*
- *Google Scholar*
- *SBCOpenLib – Simpósio Brasileiro de Informática na Educação (SBIE)*

O Simpósio Brasileiro de Informática na Educação (SBIE) foi selecionado por ser considerado uma das principais referências de publicação de informática na educação no Brasil, com índice H5 igual a 16, devido a sua contribuição em MDE e LA (Rodrigues et al., 2018) e, também, na identificação de possíveis fatores sociais e culturais que impactam a área.

A busca nas bases eletrônicas deu-se através de expressões oriundas de palavras-chave, incluindo sinônimos ou palavras relacionadas para compor os termos. Em inglês, a *string* de busca utilizada nas bases *IEEE*, *ACM* e *Google Scholar* foi ("*active learning*" OR "*active methodology*") AND ("*EDM*" OR "*education data mining*" OR "*educational data mining*") AND ("*LA*" OR "*learning analytics*") AND ("*dropout*" OR "*evasion*") AND ("*distance learning*" OR "*distance teaching*" OR "*distance education*"). Na base *Science Direct*, devido a limitação da quantidade de termos, foi ("*active learning*" OR "*active methodology*") AND ("*EDM*" OR "*educational data mining*") AND ("*LA*" OR "*learning analytics*") AND ("*dropout*") AND ("*distance learning*" OR "*distance education*"). Já na *SBIE*, a *string* de busca utilizada, em português, foi ("*aprendizagem ativa*" OU "*metodologias ativas*") E ("*MDE*" OU "*mineração de dados na educação*" OU "*mineração de dados educacionais*") E ("*abandono*" OU "*evasão*") E ("*aprendizagem à distância*" OU "*ensino a distância*" OU "*educação a distância*").

4.3 Critérios de inclusão e exclusão

Os artigos foram selecionados de acordo com os seguintes Critérios de Inclusão: estudos com foco em evasão em EaD; estudos que utilizaram técnicas de MDE e LA; estudos que abordam o uso de plataformas educacionais *on-line*; estudos que tratam sobre cursos de formação em nível médio, de graduação ou pós-graduação; estudos que discutam o uso de ferramentas para aplicação de técnicas de MDE e LA; estudos escritos em Inglês e Português; e, por fim, estudos publicados entre janeiro de 2015 e março de 2023.

Em contraponto, os Critérios de Exclusão foram: estudos com foco em evasão na educação presencial; estudos que não abordassem técnicas de MDE ou LA na evasão em EaD; estudos que discutissem o método híbrido (somente cursos 100% à distância foram considerados aceitáveis); Dissertações, Teses e livros foram excluídos.

4.4 Filtros de resultados

Para a seleção dos artigos quatro etapas foram realizadas, conforme figura 1.

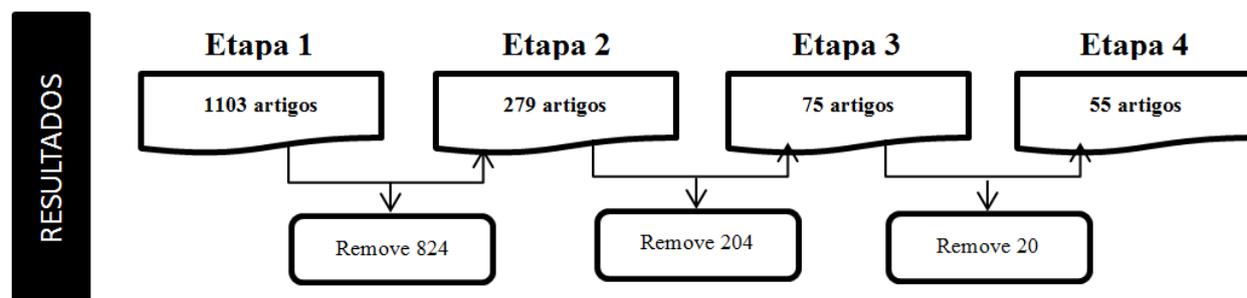


Figura 1: Resultados por etapas do processo de seleção.

A etapa 1 consistiu na pesquisa de artigos nos bancos de dados e conferências. Os artigos duplicados foram eliminados, resultando em um total de 1103 artigos reunidos. A etapa 2 compreendeu a identificação de estudos potencialmente relevantes tendo por base o título, resumo e palavras-chave, em que 824 artigos foram descartados e 279 selecionados. Na etapa 3, os estudos selecionados foram revisados com a leitura da introdução, resultados e conclusões, aplicando os critérios de inclusão e exclusão. Se a leitura dos itens propostos não era suficiente para considerar a relevância do estudo, tornava-se necessário a leitura na íntegra de tal artigo. Nesta fase, foram descartados 204 artigos e 75 selecionados. Por fim, na etapa 4, obteve-se uma lista de 55 artigos para avaliação crítica e extração de dados relevantes de cada trabalho, conforme apresenta a tabela 2.

Tabela 2: Quantidade de artigos obtidos e selecionados.

Base de dados/conferência	Artigos retornados	Artigos selecionados
IEEE ²	588	16
Science Direct ³	92	6
ACM ⁴	102	14
Google Scholar ⁵	283	9
SBIE ⁶	38	10
TOTAL	1103	55

Com a finalidade de qualificar a análise, o gráfico 1 traz o quantitativo dos artigos primários selecionados da língua inglesa e portuguesa, das bases internacionais e conferência citadas, por ano de publicação.

² IEEE - <https://www.ieee.org/index.html>.

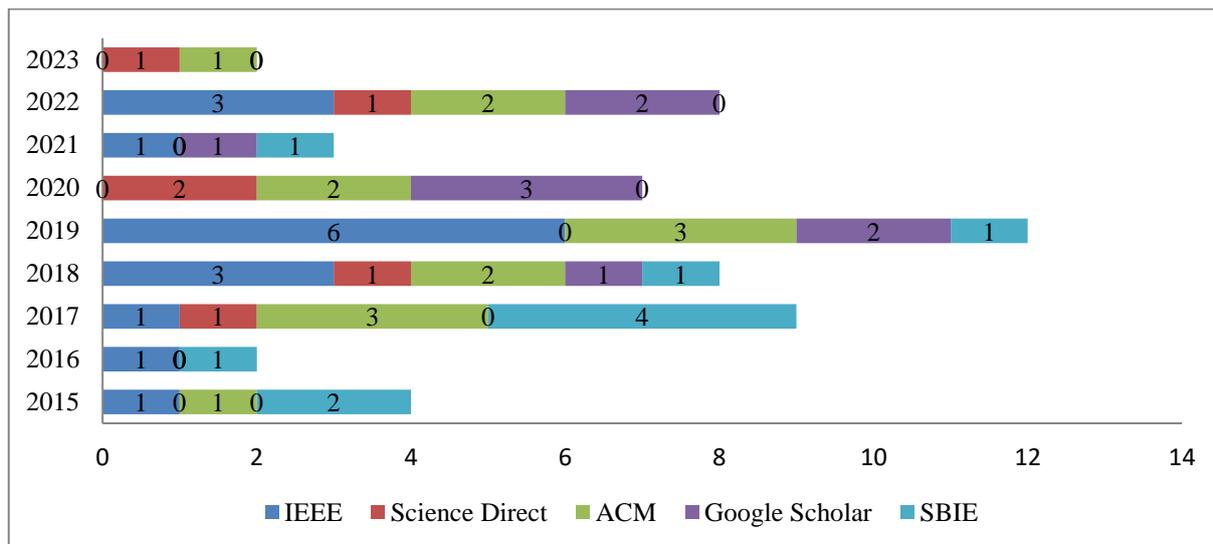
³ Elsevier Science Direct - <https://www.elsevier.com/>.

⁴ ACM Digital Library - <http://dl.acm.org/>.

⁵ Google Scholar - <https://scholar.google.com.br/>

⁶ SBIE – Simpósio Brasileiro de Informática na Educação - <https://cbie.ceie-br.org/>.

Gráfico 1: Artigos selecionados por ano de publicação.



5 Resultados e Discussões

Esta seção reúne, para cada questão de pesquisa, os resultados obtidos dos 55 artigos primários analisados, provenientes da busca por Metodologias Ativas para mitigar a evasão em EaD identificados por meio de MDE e LA.

Q1: Evasão em EaD tem sido objeto de estudo?

A evasão tem sido objeto de estudo mundial, tanto no ensino presencial quanto no ensino a distância. As técnicas de MDE e LA têm ajudado na predição, detecção, diagnóstico ou acompanhamento dos estudantes. A tabela 4 apresenta a base de dados ou conferência, o ano de publicação e autores dos 55 artigos que fundamentam a análise. Os objetivos, resultados e conclusões dos trabalhos com contribuições mais relevantes são descritos em ordem cronológica, logo após a tabela 3, de forma a responder a questão.

Tabela 3: Artigos selecionados para a pesquisa que abordam evasão em EaD.

Base de dados	Ano	Autores
IEEE	2015	Almeida Neto & Castro (2015)
	2016	Liang et al. (2016)
	2017	La Peña et al. (2017)
	2018	Isidro et al. (2018); Kostopoulos et al. (2018b); Mishra & Mishra (2018)
	2019	Brandão et al. (2019); Brito et al. (2019a); Cobos & Olmos (2019); Macedo et al. (2019); Ortigosa et al. (2019); Wang & Wang (2019)
	2021	Leite et al. (2021)
	2022	Adnan et al. (2022); Queiroga et al. (2022); Shafiq et al. (2022)
Science Direct	2017	Oeda & Hashimoto (2017)
	2018	Heidrich et al. (2018)
	2020	Tomasevic et al. (2020); Waheed et al. (2020)
	2022	Feldman-Maggor et al. (2022)
	2023	Waheed et al. (2023)

Continua na próxima página.

Tabela 3: Artigos selecionados para a pesquisa que abordam evasão em EaD. (Continuação da página anterior.)

Base de dados	Ano	Autores
ACM	2015	Kostopoulos et al. (2015)
	2017	Chen & Zhang (2017); Wang et al. (2017); Whitehill et al. (2017)
	2018	Kang & Wang (2018); Niu et al. (2018)
	2019	Borrella et al. (2019); Imran et al. (2019); Wu et al. (2019)
	2020	Islam et al. (2020); Oreshin et al. (2020)
	2022	Li et al. (2022); Ng & Lei (2022)
	2023	Tran et al. (2023)
Google Scholar	2018	Kostopoulos et al. (2018a)
	2019	Kostopoulos et al. (2019a); Kostopoulos et al. (2019b)
	2020	Goel & Goyal (2020); Karlos et al. (2020); Lemay & Doleck (2020)
	2021	Esteban et al. (2021)
	2022	Godinez & Lomibao (2022); Tamada et al. (2022)
SBIE	2015	Santos et al. (2015); Silva et al. (2015)
	2016	Santos et al. (2016)
	2017	Queiroga et al. (2017); Rabelo et al. (2017); Ramos et al. (2017); Santos & Falcão (2017)
	2018	Ramos et al. (2018)
	2019	Queiroga et al. (2019)
	2021	Nascimento et al. (2021)

A tabela 4 apresenta a relação de países, a quantidade e referências dos artigos analisados sobre o tema, identificados pela instituição de origem dos autores, o local de aplicação ou citação direta ao país no texto. Em relação ao Brasil, 8 artigos foram publicados em bases de dados internacionais e 10 no Simpósio Brasileiro de Informática na Educação (SBIE)

Tabela 4: Relação de artigos por país.

País	Quantidade	Referências
Arábia Saudita	4	Waheed et al. (2020); Islam et al. (2020); Adnan et al. (2022); Waheed et al. (2023)
Austrália	1	Tran et al. (2023)
Brasil	18	Brito et al. (2019a); Brandão et al. (2019); Almeida Neto & Castro (2015); Macedo et al. (2019); Heidrich et al. (2018); Santos et al. (2015); Silva et al. (2015); Santos et al. (2016); Ramos et al. (2017); Santos & Falcão (2017); Rabelo et al. (2017); Queiroga et al. (2017); Ramos et al. (2018); Queiroga et al. (2019); Leite et al. (2021); Nascimento et al. (2021); Queiroga et al. (2022); Tamada et al. (2022)
China	8	Liang et al. (2016); Wang & Wang (2019); Oeda & Hashimoto (2017); Chen & Zhang (2017); Niu et al. (2018); Wu et al. (2019); Ng & Lei (2022); Li et al. (2022)
Espanha	4	Ortigosa et al. (2019); Cobos & Olmos (2019); La Peña et al. (2017); Isidro et al. (2018)
Estados Unidos	5	Whitehill et al. (2017); Kang & Wang (2018); Imran et al. (2019); Borrella et al. (2019); Lemay & Doleck (2020)
Filipinas	1	Godinez & Lomibao (2022)
Grécia	6	Kostopoulos et al. (2018b); Kostopoulos et al. (2015); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Kostopoulos et al. (2019a); Karlos et al. (2020)
Índia	2	Mishra & Mishra (2018); Goel & Goyal (2020)
Israel	1	Feldman-Maggor et al. (2022)
Reino Unido	2	Shafiq et al. (2022); Esteban et al. (2021)
Rússia	1	Oreshin et al. (2020)
Sérvia	1	Tomasevic et al. (2020)
Singapura	1	Wang et al. (2017)

Q2. *Quais técnicas de MDE e LA são utilizadas para prever, detectar, diagnosticar ou monitorar a evasão em EaD?*

Em estudos recentes, várias técnicas de MDE e LA lidam com a evasão em EaD, de forma a auxiliar professores e gestores educacionais na tomada de decisão. Segundo Kostopoulos et al. (2018a), a previsão se transformou em um tópico essencial e desafiador no campo educacional, considerado um dos aspectos mais interessantes e estudados de MDE e LA.

De acordo com os artigos primários estudados, a tabela 5 apresenta as técnicas, além da quantidade e a relação dos artigos que os utilizaram. Destacam-se a “Classificação” e “Predição”, utilizadas em 48 e 31 artigos, respectivamente, no estudo de evasão em EaD, de acordo com o comportamento de aprendizagem e desempenho dos alunos na plataforma de ensino, possibilitando a obtenção de resultados positivos, intervenções oportunas e eficazes.

Tabela 5: Relação de artigos por técnica utilizada.

Técnica	Quantidade	Referências
Classificação	48	Kostopoulos et al. (2018b); Brandão et al. (2019); Ortigosa et al. (2019); Almeida Neto & Castro (2015); Liang et al. (2016); Cobos & Olmos (2019); Isidro et al. (2018); Wang & Wang (2019); Waheed et al. (2020); Heidrich et al. (2018); Tomasevic et al. (2020); Islam et al. (2020); Chen & Zhang (2017); Whitehill et al. (2017); Wang et al. (2017); Kang & Wang (2018); Niu et al. (2018); Kostopoulos et al. (2015); Wu et al. (2019); Borrella et al. (2019); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Kostopoulos et al. (2019a); Santos et al. (2015); Silva et al. (2015); Santos et al. (2016); Ramos et al. (2017); Santos & Falcão (2017); Rabelo et al. (2017); Queiroga et al. (2017); Ramos et al. (2018); Queiroga et al. (2019); Goel & Goyal (2020); Karlos et al. (2020); Lemay & Doleck (2020); Oreshin et al. (2020); Leite et al. (2021); Esteban et al. (2021); Nascimento et al. (2021); Queiroga et al. (2022); Adnan et al. (2022); Feldman-Maggor et al. (2022); Godinez & Lomibao (2022); Tamada et al. (2022); Ng & Lei (2022); Li et al. (2022); Waheed et al. (2023); Tran et al. (2023)
Predição	31	Kostopoulos et al. (2018b); Mishra & Mishra (2018); Cobos & Olmos (2019); La Peña et al. (2017); Isidro et al. (2018); Wang & Wang (2019); Waheed et al. (2020); Heidrich et al. (2018); Tomasevic et al. (2020); Whitehill et al. (2017); Wang et al. (2017); Kang & Wang (2018); Imran et al. (2019); Niu et al. (2018); Kostopoulos et al. (2015); Wu et al. (2019); Borrella et al. (2019); Kostopoulos et al. (2019b); Kostopoulos et al. (2019a); Silva et al. (2015); Santos et al. (2016); Ramos et al. (2017); Queiroga et al. (2017); Ramos et al. (2018); Queiroga et al. (2019); Lemay & Doleck (2020); Oreshin et al. (2020); Leite et al. (2021); Queiroga et al. (2022); Feldman-Maggor et al. (2022); Waheed et al. (2023)
Agrupamento	7	Brandão et al. (2019); Cobos & Olmos (2019); Macedo et al. (2019); Oeda & Hashimoto (2017); Islam et al. (2020); Santos & Falcão (2017); Shafiq et al. (2022)
Regressão	2	Kostopoulos et al. (2019b); Kostopoulos et al. (2018a)
Sumarização/ Visualização	1	Islam et al. (2020)
Não especificado	1	Brito et al. (2019a)

Vale ressaltar que em diversos artigos uma ou mais técnicas são utilizadas paralelamente. Kostopoulos et al. (2019b) utilizam “Classificação”, “Predição” e “Regressão” para prever o desempenho de estudantes através do desenvolvimento de um algoritmo de aprendizagem semi-supervisionado. Cobos & Olmos (2019) utilizam “Classificação”, “Predição” e “Agrupamento” na ferramenta desenvolvida para prever os alunos que concluirão ou não o curso.

Em concordância com Waheed et al. (2020), vários estudos implementam técnicas de Aprendizado de Máquina para analisar o comportamento e prever os alunos em risco de abandono dos estudos. No entanto, não há um consenso entre os pesquisadores sobre qual

combinação de técnicas poderia produzir melhores resultados, uma vez que a superioridade de um modelo sobre outro na previsão de evasão não pode ser afirmada, em geral, por diversos motivos, como a especificação do problema, tipo e a características de dados a serem analisados (Imran et al., 2019).

Q3. *Quais ferramentas computacionais foram utilizadas para o estudo de evasão em EaD?*

Como citado anteriormente, para o funcionamento da educação a distância, são necessárias ferramentas computacionais para disponibilizar materiais, intermediar a comunicação entre instrutores e alunos e promover o processo de ensino e aprendizagem, capaz de gerar e armazenar uma infinidade de dados significativos.

Para responder essa questão, categorizamos as respostas em 3 dimensões, conforme assuntos encontrados nos artigos estudados: (i) ferramentas utilizadas para coletar e armazenar dados; (ii) ferramentas para extrair e analisar dados armazenados; (iii) novas ferramentas desenvolvidas pelos autores dos artigos primários para o estudo da evasão.

Sobre a primeira dimensão, 47 dos 55 artigos primários divulgam as ferramentas utilizadas como fonte de dados para a coleta e armazenamento, exceto Santos & Falcão (2017), Queiroga et al. (2017), Ramos et al. (2018), Kang & Wang (2018), Brito et al. (2019a), Islam et al. (2020), Oreshin et al. (2020) e Queiroga et al. (2022), que não citam ou especificam a ferramenta utilizada para essa finalidade.

A tabela 6 apresenta os dados da primeira dimensão, ordenados pelo maior número de uso. Das 17 ferramentas relatadas, a mais citada foi o *Moodle* em 16 artigos. Demais considerações importantes das ferramentas: 5 MOOCs diferentes foram utilizados, o que representa uma grande aceitação dessa plataforma no processo de ensino e aprendizagem a distância; Wang & Wang (2019), Oeda & Hashimoto (2017), Ng & Lei (2022) e Li et al. (2022) não são específicos na identificação de qual MOOC utilizam; Imran et al. (2019) utilizam dois diferentes MOOCs; Ortigosa et al. (2019), Heidrich et al. (2018) e Shafiq et al. (2022) utilizam duas ferramentas.

Tabela 6: Ferramentas utilizadas nos artigos para coleta e armazenamento de dados.

Ferramentas	Quantidade	Referências
Moodle	16	Brandão et al. (2019); Ortigosa et al. (2019); La Peña et al. (2017); Macedo et al. (2019); Heidrich et al. (2018); Santos et al. (2015); Silva et al. (2015); Santos et al. (2016); Ramos et al. (2017); Rabelo et al. (2017); Queiroga et al. (2019); Nascimento et al. (2021); Feldman-Maggor et al. (2022); Tamada et al. (2022); Tran et al. (2023); Waheed et al. (2023)
<i>Hellenic Open University</i> (HOU)	5	Kostopoulos et al. (2018b); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Kostopoulos et al. (2019a); Karlos et al. (2020)
<i>Open University Learning Analytics</i> (OULA)	5	Waheed et al. (2020); Tomasevic et al. (2020); Esteban et al. (2021); Shafiq et al. (2022); Adnan et al. (2022)
MOOC	4	Wang & Wang (2019); Oeda & Hashimoto (2017); Ng & Lei (2022); Li et al. (2022)
MOOC da EdX	4	Cobos & Olmos (2019); Isidro et al. (2018); Chen & Zhang (2017); Lemay & Doleck (2020)
MOOC XuetangX	3	Liang et al. (2016); Wang et al. (2017); Goel & Goyal (2020)
MOOC da MitX	2	Imran et al. (2019); Borrella et al. (2019)
MOOC da HarvardX	2	Whitehill et al. (2017); Imran et al. (2019)
<i>Index of Learning Style Questionnaire</i> (ILSQ)	1	Heidrich et al. (2018)
<i>Universitas-XXI</i>	1	Ortigosa et al. (2019)
ColabWeb	1	Almeida Neto & Castro (2015)

Continua na próxima página.

Tabela 6: Ferramentas utilizadas nos artigos para coleta e armazenamento de dados. (Continuação da página anterior)

Ferramentas	Quantidade	Referências
ICourse163	1	Niu et al. (2018)
KDDCup	1	Wu et al. (2019)
Oracle	1	Mishra & Mishra (2018)
Kell	1	Kostopoulos et al. (2015)
Jupyter	1	Shafiq et al. (2022)
PyCaret	1	Leite et al. (2021)

Sobre a segunda dimensão, foram utilizadas 21 ferramentas para extrair ou analisar os dados armazenados oriundos da plataforma de ensino, conforme a tabela 7. Destaque para a ferramenta *Waikato Environment for Knowledge Analysis (Weka)*, utilizada em 11 trabalhos para essa finalidade, uma importante aliada no combate a evasão em EaD. Um artigo pode utilizar mais de uma ferramenta, assim como em Macedo et al. (2019), Brito et al. (2019a), Islam et al. (2020), Leite et al. (2021) e Shafiq et al. (2022).

Tabela 7: Ferramentas utilizadas nos artigos para extrair ou analisar dados.

Ferramentas	Quantidade	Referências
Weka	11	Mishra & Mishra (2018); Cobos & Olmos (2019); Kostopoulos et al. (2019b); Santos et al. (2015); Silva et al. (2015); Santos et al. (2016); Santos & Falcão (2017); Rabelo et al. (2017); Lemay & Doleck (2020); Esteban et al. (2021); Tran et al. (2023)
R Studio	4	Macedo et al. (2019); Islam et al. (2020); Ramos et al. (2018); Godinez & Lomibao (2022)
MatLab	2	Mishra & Mishra (2018); Tomasevic et al. (2020)
JCLAL (<i>Java Class Library for Active Learning</i>)	1	Kostopoulos et al. (2018b)
<i>DB Extractor</i>	1	Brito et al. (2019a)
<i>Google Charts</i>	1	Brito et al. (2019a)
ERP Oracle	1	Ortigosa et al. (2019)
SPMF	1	Almeida Neto & Castro (2015)
MySQL WorkBench	1	Macedo et al. (2019)
PGAdmin III	1	Macedo et al. (2019)
Microsoft Excel	1	Macedo et al. (2019)
MapReduce	1	Islam et al. (2020)
Hadoop	1	Islam et al. (2020)
SPSS	1	Santos et al. (2015)
R Project	1	Ramos et al. (2018)
Plotly	1	Shafiq et al. (2022)
Matplotlib	1	Shafiq et al. (2022)
Seaborn	1	Shafiq et al. (2022)
Numpy	1	Leite et al. (2021)
Microsoft Power BI	1	Leite et al. (2021)
Optuna	1	Ng & Lei (2022)

Na terceira dimensão, destacam-se 11 novos produtos, conforme a tabela 8. A ferramenta apresentada em Cobos & Olmos (2019), EdX-MAS+, foi testada em 7 MOOCs e utiliza 3 diferentes técnicas para analisar os dados contidos nas plataformas de ensino a distância: classificação, predição e agrupamento. Brito et al. (2019a) apresentam a ferramenta *Dropout Risk Report*, capaz de demonstrar uma lista de alunos em risco de evasão por meio de relatórios gráficos. Chen & Zhang (2017), Whitehill et al. (2017), Santos & Falcão (2017), Queiroga et al. (2017) e Adnan et al. (2022) apresentam os resultados obtidos pelas ferramentas desenvolvidas, no entanto, não divulgam os nomes de seus sistemas inovadores.

Tabela 8: Ferramentas desenvolvidas.

Ferramentas	Técnicas de MDE/LA	Referências
SPA (Sistema de Predicção de Abandono)	Classificação	Ortigosa et al. (2019)
<i>Dropout Risk Report</i>	Não especificado*	Brito et al. (2019a)
EdX-MAS+ (<i>Model Analyzer System Plus</i>)	Classificação, Predição, Agrupamento	Cobos & Olmos (2019)
LOGIT_Act	Predição	La Peña et al. (2017)
<i>Prediction Tool</i>	Classificação, Regressão	Kostopoulos et al. (2018a)
REA 2.0 (Roda de Estados de Ânimo)	Classificação	Santos et al. (2015)
Nomes não divulgados	Classificação, Predição, Agrupamento	Chen & Zhang (2017), Whitehill et al. (2017), Santos & Falcão (2017), Queiroga et al. (2017), Adnan et al. (2022)

Q4. Quais algoritmos foram utilizados para o estudo de evasão em EaD?

Diversos são os algoritmos existentes advindos da Mineração de Dados, cada qual com a sua finalidade de aplicação. Tratando-se de predição, detecção, diagnóstico ou acompanhamento de alunos aptos a evadir, muitos têm sido utilizados nos meios educacionais.

A tabela 9 apresenta 79 algoritmos encontrados nos artigos primários, com destaque para *Random Forest*, *Logistic Regression*, *SVM*, *Naive Bayes* e *AdaBoost*, utilizados em 25, 22, 20, 17 e 12, respectivamente. O algoritmo *Random Forest* tem uma boa aceitação, pois é simples de implementar e tem a metodologia de funcionamento baseada na criação de árvores de decisão, capaz de ajudar nas tarefas de classificação, predição e regressão.

Tabela 9: Algoritmos utilizados nos artigos.

Algoritmo	Quantidade	Referências
<i>Random Forest</i>	25	Kostopoulos et al. (2018b); Brandão et al. (2019); Ortigosa et al. (2019); Liang et al. (2016); Mishra & Mishra (2018); Cobos & Olmos (2019); Wang & Wang (2019); Chen & Zhang (2017); Wang et al. (2017); Kang & Wang (2018); Wu et al. (2019); Borrella et al. (2019); Kostopoulos et al. (2019b); Kostopoulos et al. (2019a); Santos et al. (2016); Santos & Falcão (2017); Queiroga et al. (2017); Queiroga et al. (2019); Oreshin et al. (2020); Leite et al. (2021); Esteban et al. (2021); Queiroga et al. (2022); Shafiq et al. (2022); Adnan et al. (2022); Tamada et al. (2022)
<i>Logistic Regression</i>	22	Kostopoulos et al. (2018b); Liang et al. (2016); La Peña et al. (2017); Wang & Wang (2019); Waheed et al. (2020); Tomasevic et al. (2020); Wang et al. (2017); Kang & Wang (2018); Borrella et al. (2019); Ramos et al. (2017); Ramos et al. (2018); Queiroga et al. (2019); Lemay & Doleck (2020); Oreshin et al. (2020); Leite et al. (2021); Esteban et al. (2021); Queiroga et al. (2022); Shafiq et al. (2022); Adnan et al. (2022); Feldman-Maggor et al. (2022); Tamada et al. (2022); Waheed et al. (2023)
<i>Support Vector Machine (SVM)</i>	20	Liang et al. (2016); Cobos & Olmos (2019); La Peña et al. (2017); Isidro et al. (2018); Wang & Wang (2019); Waheed et al. (2020); Heidrich et al. (2018); Tomasevic et al. (2020); Wang et al. (2017); Kang & Wang (2018); Wu et al. (2019); Santos et al. (2016); Ramos et al. (2017); Ramos et al. (2018); Leite et al. (2021); Esteban et al. (2021); Adnan et al. (2022); Tamada et al. (2022); Tran et al. (2023); Waheed et al. (2023)

Continua na próxima página.

Tabela 9: Algoritmos utilizados nos artigos. (Continuação da página anterior)

Algoritmo	Quantidade	Referências
<i>Naive Bayes</i>	17	Kostopoulos et al. (2018b); Cobos & Olmos (2019); Isidro et al. (2018); Heidrich et al. (2018); Tomasevic et al. (2020); Kang & Wang (2018); Kostopoulos et al. (2015); Kostopoulos et al. (2018a); Silva et al. (2015); Queiroga et al. (2019); Lemay & Doleck (2020); Leite et al. (2021); Esteban et al. (2021); Queiroga et al. (2022); Adnan et al. (2022); Godinez & Lomibao (2022); Tamada et al. (2022)
<i>AdaBoost</i>	12	Brandão et al. (2019); Isidro et al. (2018); Wang et al. (2017); Wu et al. (2019); Kostopoulos et al. (2018a); Santos et al. (2016); Queiroga et al. (2019); Leite et al. (2021); Esteban et al. (2021); Queiroga et al. (2022); Adnan et al. (2022); Tran et al. (2023)
<i>Decision Tree</i>	10	Wang & Wang (2019); Tomasevic et al. (2020); Wu et al. (2019); Leite et al. (2021); Esteban et al. (2021); Shafiq et al. (2022); Adnan et al. (2022); Feldman-Maggor et al. (2022); Tamada et al. (2022); Waheed et al. (2023)
<i>J48 Decision Tree</i>	10	Kostopoulos et al. (2018b); Heidrich et al. (2018); Santos et al. (2015); Silva et al. (2015); Santos et al. (2016); Rabelo et al. (2017); Queiroga et al. (2017); Queiroga et al. (2019); Lemay & Doleck (2020); Esteban et al. (2021)
<i>K-nearest Neighbor (KNN)</i>	10	Cobos & Olmos (2019); Tomasevic et al. (2020); Kang & Wang (2018); Kostopoulos et al. (2019b); Kostopoulos et al. (2019a); Ramos et al. (2017); Ramos et al. (2018); Karlos et al. (2020); Adnan et al. (2022); Waheed et al. (2023)
<i>Multilayer Perceptrons</i>	8	Kostopoulos et al. (2018b); Santos et al. (2016); Queiroga et al. (2017); Queiroga et al. (2019); Leite et al. (2021); Esteban et al. (2021); Queiroga et al. (2022); Tran et al. (2023)
<i>Gradient Boosting Decision Tree (GBDT)</i>	7	Liang et al. (2016); Cobos & Olmos (2019); Wang & Wang (2019); Wang et al. (2017); Wu et al. (2019); Tamada et al. (2022); Waheed et al. (2023)
<i>Linear Regression</i>	6	Tomasevic et al. (2020); Kang & Wang (2018); Wu et al. (2019); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Karlos et al. (2020)
<i>Artificial Neural Network (ANN)</i>	6	Heidrich et al. (2018); Tomasevic et al. (2020); Ramos et al. (2017); Ramos et al. (2018); Nascimento et al. (2021); Waheed et al. (2023)
<i>K-means</i>	6	Brandão et al. (2019); Oeda & Hashimoto (2017); Islam et al. (2020); Nascimento et al. (2021); Shafiq et al. (2022); Tamada et al. (2022)
<i>Sequential Minimal Optimization (SMO)</i>	5	Kostopoulos et al. (2018b); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Lemay & Doleck (2020); Esteban et al. (2021)
<i>Bayes Net</i>	4	Kostopoulos et al. (2018b); Silva et al. (2015); Santos et al. (2016); Queiroga et al. (2017)
<i>Classification and Regression Trees (CART)</i>	4	Brandão et al. (2019); Mishra & Mishra (2018); Isidro et al. (2018); Silva et al. (2015)
<i>C4.5</i>	4	Mishra & Mishra (2018); Kostopoulos et al. (2015); Kostopoulos et al. (2018a); Tran et al. (2023)
<i>Extra Tree Classifier (EXTRA)</i>	3	Kostopoulos et al. (2018a); Leite et al. (2021); Adnan et al. (2022)
<i>Gaussian Naive Bayes</i>	3	Wang et al. (2017); Wu et al. (2019); Kostopoulos et al. (2019a)
<i>Long Short Term Memory (LSTM)</i>	3	Isidro et al. (2018); Wu et al. (2019); Waheed et al. (2023)
<i>Support Vector Machine (SVM) + RBF Kernel</i>	3	Wang et al. (2017); Wu et al. (2019); Leite et al. (2021)

Continua na próxima página.

Tabela 9: Algoritmos utilizados nos artigos. (Continuação da página anterior)

Algoritmo	Quantidade	Referências
<i>Stochastic Gradient Descent (SGD)</i>	3	Goel & Goyal (2020); Esteban et al. (2021); Adnan et al. (2022)
<i>Deep Artificial Neural Network (Deep ANN)</i>	2	Waheed et al. (2020); Imran et al. (2019)
<i>Extreme Gradient Boosting (XGBoost)</i>	2	Niu et al. (2018); Oreshin et al. (2020)
<i>M5' Rules</i>	2	Kostopoulos et al. (2019b); Kostopoulos et al. (2018a)
<i>M5 Model Tree</i>	2	Kostopoulos et al. (2019b); Kostopoulos et al. (2018a)
<i>Radial Basis Function (RBF)</i>	2	Kostopoulos et al. (2018a); Esteban et al. (2021)
<i>Stochastic Gradient Boosting (SGB)</i>	2	Cobos & Olmos (2019); Adnan et al. (2022)
<i>JRip</i>	2	Santos et al. (2016); Lemay & Doleck (2020)
<i>IBK</i>	2	Santos et al. (2016); Lemay & Doleck (2020)
<i>Light Gradient Boosting Model (LightGBM)</i>	2	Leite et al. (2021); Ng & Lei (2022)
<i>Categorical Boosting (CatBoost)</i>	2	Oreshin et al. (2020); Leite et al. (2021)
<i>Bagging</i>	2	Esteban et al. (2021); Adnan et al. (2022)
<i>Bayesian Regression</i>	1	Tomasevic et al. (2020)
<i>LogitBoost</i>	1	Kostopoulos et al. (2018a)
<i>Simple Logistic</i>	1	Queiroga et al. (2017)
<i>K-means++</i>	1	Oeda & Hashimoto (2017)
<i>K-mediods</i>	1	Oeda & Hashimoto (2017)
<i>C5.0</i>	1	Ortigosa et al. (2019)
<i>Rotation Forest</i>	1	Kostopoulos et al. (2018a)
<i>Apriori Inverse</i>	1	Almeida Neto & Castro (2015)
<i>Boosted Logistic Regression</i>	1	Cobos & Olmos (2019)
<i>Gradient Boosting Regressor (GBR)</i>	1	Karlos et al. (2020)
<i>Neuronal Network (NN)</i>	1	Cobos & Olmos (2019)
<i>Convolutional Neural Network (CNN)</i>	1	Wu et al. (2019)
<i>Bayesian Generalized Linear Model (BGLN)</i>	1	Cobos & Olmos (2019)
<i>Feed-Forward Neural Network</i>	1	La Peña et al. (2017)
<i>Probabilistic Ensemble Simplified Fuzzy (PESFAM)</i>	1	La Peña et al. (2017)
<i>Fuzzy C-means</i>	1	Macedo et al. (2019)
<i>Dynamic Time Warping</i>	1	Oeda & Hashimoto (2017)
<i>ID3</i>	1	Rabelo et al. (2017)
<i>Inter-Quartile Range</i>	1	Islam et al. (2020)
<i>L2-Regularized Logistic Regression</i>	1	Whitehill et al. (2017)
<i>Reduced Error Pruning Tree (REPTree)</i>	1	Kostopoulos et al. (2018a)
<i>BFTree</i>	1	Silva et al. (2015)
<i>3-Nearest Neighbor</i>	1	Kostopoulos et al. (2018a)
<i>Expectation Maximization (EM)</i>	1	Santos & Falcão (2017)
<i>Generalized Predictive Control (GPC)</i>	1	Leite et al. (2021)
<i>Gradient Bi-Conjugate (GBC)</i>	1	Leite et al. (2021)
<i>Ridge</i>	1	Leite et al. (2021)
<i>Boruta</i>	1	Godinez & Lomibao (2022)
<i>Random Tree</i>	1	Esteban et al. (2021)
<i>Decision Stump</i>	1	Esteban et al. (2021)

Continua na próxima página.

Tabela 9: Algoritmos utilizados nos artigos. (Continuação da página anterior)

Algoritmo	Quantidade	Referências
<i>NNge</i>	1	Esteban et al. (2021)
<i>ZeroR</i>	1	Esteban et al. (2021)
<i>OneR</i>	1	Esteban et al. (2021)
<i>PART</i>	1	Esteban et al. (2021)
<i>Ridor</i>	1	Esteban et al. (2021)
<i>RMSProp</i>	1	Goel & Goyal (2020)
<i>Adagrad</i>	1	Goel & Goyal (2020)
<i>Adadelta</i>	1	Goel & Goyal (2020)
<i>Adam</i>	1	Goel & Goyal (2020)
<i>Adamax</i>	1	Goel & Goyal (2020)
<i>Nadam</i>	1	Goel & Goyal (2020)
<i>Weka Deep Learning 4J</i>	1	Lemay & Doleck (2020)
<i>Bayesian Probabilistic Neural Network (Bayesian PNN)</i>	1	Tran et al. (2023)
<i>Voting</i>	1	Adnan et al. (2022)
<i>ResNet</i>	1	Li et al. (2022)
<i>InceptionTime</i>	1	Li et al. (2022)
Não Informado	2	Brito et al. (2019a); Imran et al. (2019)

Em relação aos 79 algoritmos citados, verificam-se autores que utilizam um ou mais algoritmos para a obtenção dos resultados. Esteban et al. (2021), Leite et al. (2021) e Kostopoulos et al. (2018a), por exemplo, utilizam 18, 14 e 12 algoritmos, respectivamente, com o objetivo de prever o desempenho dos alunos. Almeida Neto & Castro (2015), Macedo et al. (2019), Chen & Zhang (2017), Whitehill et al. (2017), Niu et al. (2018), Ng & Lei (2022) e Santos et al. (2015) utilizam apenas 1 algoritmo, e Brito et al. (2019a) e Imran et al. (2019) não citam os algoritmos utilizados.

Com o propósito de inovação, novos algoritmos foram desenvolvidos na tentativa de contribuir com a identificação e mitigação da evasão em EaD. A tabela 10 apresenta 13 algoritmos de MDE e LA. Kostopoulos et al. (2018a), por exemplo, utilizam 12 algoritmos e propõem o desenvolvimento de um, embora não atribuído o nome. O Algoritmo *MSSRA*, proposto por Kostopoulos et al (2019b), utiliza as técnicas de classificação, predição e regressão para auxiliar na identificação de possíveis evasões dos cursos ofertados.

Tabela 10: Algoritmos desenvolvidos.

Algoritmo	Técnicas de MDE/LA	Referências
<i>PCA - Principal Component Analysis</i>	Predição	Mishra & Mishra (2018)
<i>E-LSTM</i>	Classificação, Predição	Wang & Wang (2019)
<i>ConRec Network</i>	Classificação, Predição	Wang et al. (2017)
<i>CLMS-Net</i>	Classificação, Predição	Wu et al. (2019)
<i>MSSRA</i>	Classificação, Predição, Regressão	Kostopoulos et al. (2019b)
<i>Extra</i>	Classificação, Predição	Kostopoulos et al. (2019a)
<i>Gradient Bi-Conjugate (GBC)</i>	Classificação, Predição	Kostopoulos et al. (2019a)
<i>CRISP-EDM</i>	Classificação, Predição	Ramos et al. (2017)
<i>Gaussian-Bernoulli mixed Naïve Bayes</i>	Classificação	Godinez & Lomibao (2022)
Nomes não divulgados	Classificação, Predição, Regressão	Kostopoulos et al. (2018a); Queiroga et al. (2019); Queiroga et al. (2022); Li et al. (2022)

Não há como distinguir, entre os algoritmos desenvolvidos, aquele que teve o melhor desempenho e resultado na aplicação. O fato é que, entre 2015 e 2023, novas propostas para minerar dados educacionais e auxiliar a análise de aprendizagem foram realizadas, o que

demonstra mais uma vez a relevância científica da pesquisa e a busca por novas técnicas e soluções para mitigar a evasão.

Q5. *Quais atributos foram utilizados no estudo de evasão em EaD?*

Diversas são as ferramentas computacionais utilizadas na educação a distância que coleta dados acadêmicos dos usuários cadastrados, seja o professor, tutor, aluno ou gestor. Esses dados são armazenados em colunas do banco de dados das plataformas de ensino, denominadas atributos, associando a eles um valor, ou seja, uma característica do dado.

Partindo dessa definição e para responder essa questão, classifica-se as respostas em 4 características de dados: demográficos, comportamentais, interação e desempenho. Como metodologia, os atributos foram coletados de cada artigo, sendo que 42 trabalham com análise de dados de interação, 39 com dados de desempenho, 34 com dados comportamentais e 20 com dados demográficos.

Vale ressaltar que um artigo pode ter utilizado um ou mais de um tipo de categoria de dados, como, por exemplo, demográfico + interação, interação + desempenho, comportamento + interação, ou demográfico + interação + desempenho. A tabela 11 apresenta a relação de artigos e tipo de dados utilizados em cada um.

Tabela 11: Relação de artigos por tipo de dados.

Referências	Tipo de Dados			
	Demográficos	Comportamental	Interações	Desempenho
Adnan et al. (2022)	X	X		X
Almeida Neto & Castro (2015)		X	X	X
Borrella et al. (2019)			X	
Brandão et al. (2019)			X	X
Brito et al. (2019a)		X	X	X
Chen & Zhang (2017)		X	X	X
Cobos & Olmos (2019)			X	
Esteban et al. (2021)				X
Feldman-Maggor et al. (2022)	X	X	X	X
Godinez & Lomibao (2022)	X			X
Goel & Goyal (2020)	X	X		
Heidrich et al. (2018)		X	X	X
Imran et al. (2019)		X	X	
Isidro et al. (2018)	X		X	X
Islam et al. (2020)			X	X
Kang & Wang (2018)	X	X		X
Karlos et al. (2020)	X	X	X	X
Kostopoulos et al. (2015)	X	X	X	X
Kostopoulos et al. (2018a)	X		X	X
Kostopoulos et al. (2018b)	X		X	X
Kostopoulos et al. (2019a)	X	X	X	X
Kostopoulos et al. (2019b)	X	X	X	X
La Peña et al. (2017)				X
Leite et al. (2021)		X		
Lemay & Doleck (2020)			X	
Li et al. (2022)		X		
Liang et al. (2016)		X	X	
Macedo et al. (2019)			X	X
Mishra & Mishra (2018)		X	X	X
Nascimento et al. (2021)		X	X	X
Ng & Lei (2022)		X		
Niu et al. (2018)	X	X	X	X

Continua na próxima página.

Tabela 11: Relação de artigos por tipo de dados. (Continuação da página anterior)

Referências	Tipo de Dados			
	Demográficos	Comportamental	Interações	Desempenho
Oeda & Hashimoto (2017)			X	
Oreshin et al. (2020)	X	X		X
Ortigosa et al. (2019)	X		X	X
Queiroga et al. (2017)			X	
Queiroga et al. (2019)			X	
Queiroga et al. (2022)		X		
Rabelo et al. (2017)		X	X	X
Ramos et al. (2017)		X	X	
Ramos et al. (2018)		X	X	X
Santos et al. (2015)		X	X	X
Santos et al. (2016)		X	X	X
Santos & Falcão (2017)			X	X
Shafiq et al. (2022)	X	X		X
Silva et al. (2015)			X	
Tamada et al. (2022)	X	X	X	X
Tomasevic et al. (2020)	X	X	X	X
Tran et al. (2023)			X	X
Waheed et al. (2020)	X		X	X
Waheed et al. (2023)	X	X		X
Wang & Wang (2019)		X	X	
Wang et al. (2017)		X	X	X
Whitehill et al. (2017)			X	X
Wu et al. (2019)		X	X	X
TOTAL	20	34	42	39

Os trabalhos de Feldman-Maggor et al. (2022), Karlos et al. (2020), Kostopoulos et al. (2015), Kostopoulos et al. (2019a), Kostopoulos et al. (2019b), Niu et al. (2018), Tamada et al. (2022) e Tomasevic et al. (2020) utilizam as 4 características de dados.

Os dados demográficos são atributos de definição de usuários das plataformas de ensino armazenados como registros administrativos, pré ou pós-ingresso, nos cursos ofertados. A tabela 12 apresenta a relação de 12 atributos que caracterizam esses dados, a quantidade de artigos que os utilizam e as referências. Dos atributos relatados, “Sexo” é o mais utilizado (15 artigos), seguido pela “Idade” (12 artigos). Isidro et al. (2018), Waheed et al. (2020), Niu et al. (2018) e Waheed et al. (2023) utilizam dados demográficos para a análise, no entanto, não são específicos na definição de quais atributos.

Tabela 12: Relação de artigos por atributo de dados demográficos.

Atributo	Quantidade	Referências
Sexo	15	Kostopoulos et al. (2018b); Ortigosa et al. (2019); Tomasevic et al. (2020); Kang & Wang (2018); Kostopoulos et al. (2015); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Kostopoulos et al. (2019a); Shafiq et al. (2022); Adnan et al. (2022); Feldman-Maggor et al. (2022); Godinez & Lomibao (2022); Goel & Goyal (2020); Karlos et al. (2020); Tamada et al. (2022)
Idade	12	Kostopoulos et al. (2018b); Ortigosa et al. (2019); Tomasevic et al. (2020); Kang & Wang (2018); Kostopoulos et al. (2015); Kostopoulos et al. (2018a); Shafiq et al. (2022); Adnan et al. (2022); Feldman-Maggor et al. (2022); Godinez & Lomibao (2022); Goel & Goyal (2020); Tamada et al. (2022)
Região/Cidade	4	Tomasevic et al. (2020); Kang & Wang (2018); Shafiq et al. (2022); Adnan et al. (2022)

Continua na próxima página.

Tabela 12: Relação de artigos por atributo de dados demográficos. (Continuação da página anterior.)

Atributo	Quantidade	Referências
Número de filhos	3	Kostopoulos et al. (2018b); Kostopoulos et al. (2015); Kostopoulos et al. (2018a)
Conhecimento de informática	3	Kostopoulos et al. (2018b); Kostopoulos et al. (2018a); Godinez & Lomibao (2022)
Estado civil	2	Kostopoulos et al. (2018b); Kostopoulos et al. (2018a)
Uso de computador	2	Kostopoulos et al. (2018b); Kostopoulos et al. (2018a)
Tipo de ocupação	2	Kostopoulos et al. (2015); Kostopoulos et al. (2018a)
Nacionalidade	2	Kostopoulos et al. (2015); Oreshin et al. (2020)
Quantidade de membro familiar	2	Godinez & Lomibao (2022); Tamada et al. (2022)
Horário de trabalho	1	Kostopoulos et al. (2018b)
Renda familiar	1	Tamada et al. (2022)

Os dados comportamentais representam os mecanismos utilizados pelo usuário ao acessar a plataforma educacional. A tabela 13 apresenta 3 atributos extraídos dos registros de *logs* das plataformas de ensino, a quantidade de artigos e as referências. De acordo com os dados extraídos, 31 artigos estudam o atributo “Acesso ao curso”, considerando que a quantidade de vezes, a data e a hora em que o usuário faz o *login* na plataforma são importantes para verificar a sua vontade, disponibilidade e a motivação para a aprendizagem. Em relação ao atributo “Duração do *login*”, 8 artigos utilizam esses dados para verificar quanto tempo o usuário permanece na plataforma de ensino, desde a sua entrada até a saída do ambiente. Por fim, 5 artigos citam os dados de “Navegadores web” para o acesso a plataforma de ensino como um importante registro, capaz de auxiliar na identificação de possíveis falhas no processo de interação do usuário com o ambiente de aprendizagem.

Tabela 13: Relação de artigos por atributo de dados comportamentais.

Atributo	Quantidade	Referências
Acesso ao curso	31	Brito et al. (2019a); Almeida Neto & Castro (2015); Liang et al. (2016); Mishra & Mishra (2018); Heidrich et al. (2018); Tomasevic et al. (2020); Chen & Zhang (2017); Wang et al. (2017); Imran et al. (2019); Niu et al. (2018); Kostopoulos et al. (2015); Wu et al. (2019); Kostopoulos et al. (2019b); Kostopoulos et al. (2019a); Santos et al. (2015); Santos et al. (2016); Ramos et al. (2017); Rabelo et al. (2017); Ramos et al. (2018); Queiroga et al. (2022); Shafiq et al. (2022); Leite et al. (2021); Adnan et al. (2022); Feldman-Maggor et al. (2022); Waheed et al. (2023); Goel & Goyal (2020); Karlos et al. (2020); Tamada et al. (2022); Nascimento et al. (2021); Ng & Lei (2022); Li et al. (2022)
Duração de <i>login</i>	8	Wang & Wang (2019); Kang & Wang (2018); Kostopoulos et al. (2015); Wu et al. (2019); Kostopoulos et al. (2019a); Ramos et al. (2017); Ramos et al. (2018); Nascimento et al. (2021)
Navegadores	5	Liang et al. (2016); Mishra & Mishra (2018); Wang & Wang (2019); Wang et al. (2017); Kostopoulos et al. (2015)

Em relação aos dados de interação, que representam o envolvimento do usuário com as funcionalidades da plataforma educacional *on-line* após o *login*, os artigos primários citam 12 atributos. A tabela 14 apresenta 27 artigos que utilizam o atributo “Fórum de discussão” como objeto de estudo, uma vez que a não participação dos alunos nessa atividade ou o que ele escreve representa sinais de desânimo ou falta de motivação no curso e, conseqüentemente, a evasão. Em seguida, 16 artigos citam o atributo “Vídeo”, que representa o comportamento do usuário ao assistir uma vídeo-aula como, por exemplo, quantas vezes ele pausa, reproduz e inicia a exibição. Outro atributo citado e bastante relevante é o “Mensageiro”, que permite a comunicação assíncrona entre professor, tutor e aluno.

Tabela 14: Relação de artigos por atributo de dados de interação.

Atributo	Quantidade	Referências
Fórum de discussão	27	Brito et al. (2019a); Brandão et al. (2019); Ortigosa et al. (2019); Almeida Neto & Castro (2015); Liang et al. (2016); Mishra & Mishra (2018); Cobos & Olmos (2019); Macedo et al. (2019); Isidro et al. (2018); Wang & Wang (2019); Heidrich et al. (2018); Chen & Zhang (2017); Wang et al. (2017); Imran et al. (2019); Niu et al. (2018); Kostopoulos et al. (2015); Kostopoulos et al. (2019b); Kostopoulos et al. (2019a); Santos et al. (2015); Silva et al. (2015); Santos et al. (2016); Santos & Falcão (2017); Rabelo et al. (2017); Ramos et al. (2018); Karlos et al. (2020); Nascimento et al. (2021); Tamada et al. (2022)
Vídeos	16	Liang et al. (2016); Mishra & Mishra (2018); Cobos & Olmos (2019); Macedo et al. (2019); Wang & Wang (2019); Chen & Zhang (2017); Whitehill et al. (2017); Wang et al. (2017); Imran et al. (2019); Niu et al. (2018); Kostopoulos et al. (2015); Wu et al. (2019); Santos et al. (2016); Santos & Falcão (2017); Lemay & Doleck (2020); Feldman-Maggor et al. (2022)
Mensageiro	12	Kostopoulos et al. (2018b); Brandão et al. (2019); Ortigosa et al. (2019); Almeida Neto & Castro (2015); Macedo et al. (2019); Kostopoulos et al. (2015); Kostopoulos; et al. (2019b); Kostopoulos et al. (2018a); Kostopoulos et al. (2019a); Ramos et al. (2017); Ramos et al. (2018); Tamada et al. (2022)
Materiais disponibilizados	7	Brito et al. (2019a); Macedo et al. (2019); Wang & Wang (2019); Heidrich et al. (2018); Santos et al. (2015); Santos et al. (2016); Santos & Falcão (2017)
Wiki	4	Mishra & Mishra (2018); Wang & Wang (2019); Kostopoulos et al. (2015); Santos et al. (2016)
Bate papo	4	Brandão et al. (2019); Heidrich et al. (2018); Santos et al. (2016); Tamada et al. (2022)
Páginas visitadas	3	Liang et al. (2016); Macedo et al. (2019); Chen and Zhang (2017)
Web-conferências	2	Ramos et al. (2017); Ramos et al. (2018)
Notícias	1	Kostopoulos et al. (2019b)
E-mail	1	Borrella et al. (2019)

Os artigos de Oeda & Hashimoto (2017), Waheed et al. (2020), Tomasevic et al. (2020), Islam et al. (2020), Queiroga et al. (2017), Queiroga et al. (2019) e Tran et al. (2023) relatam a utilização de dados de interação para a análise, no entanto, não foram específicos na definição de quais atributos são utilizados.

Tabela 15: Relação de artigos por atributo de dados de desempenho.

Atributo	Quantidade	Referências
Tarefas	19	Brito et al. (2019a); Ortigosa et al. (2019); Almeida Neto & Castro (2015); Heidrich et al. (2018); Wang et al. (2017); Kostopoulos et al. (2015); Wu et al. (2019); Santos et al. (2015); Santos et al. (2016); Santos & Falcão (2017); Rabelo et al. (2017); Ramos et al. (2018); Oreshin et al. (2020); Esteban et al. (2021); Shafiq et al. (2022); Adnan et al. (2022); Feldman-Maggor et al. (2022); Godinez & Lomibao (2022); Tamada et al. (2022)
Questionários	16	Brito et al. (2019a); Brandão et al. (2019); Ortigosa et al. (2019); Mishra & Mishra (2018); Macedo et al. (2019); Chen & Zhang (2017); Whitehill et al. (2017); Niu et al. (2018); Santos et al. (2016); Santos & Falcão (2017); Rabelo et al. (2017); Nascimento et al. (2021); Shafiq et al. (2022); Adnan et al. (2022); Tamada et al. (2022); Tran et al. (2023)
Avaliação escrita	5	Kostopoulos et al. (2018b); La Peña et al. (2017); Niu et al. (2018); Kostopoulos et al. (2015); Nascimento et al. (2021)

Como última análise de atributos utilizados nos artigos primários, têm-se os dados de desempenho oriundos de 3 atributos citados que representam o grau de aprendizado e notas

obtidas no processo. Os dados da tabela 15 apresentam os atributos que indicam o nível de aprendizagem dos alunos. O atributo “Tarefas” foi utilizado para análise em 19 artigos e armazena a nota obtida pelos alunos nos exercícios e atividades avaliativas preparadas pelos professores no ambiente virtual. O atributo “Questionários”, utilizado em 16 artigos, armazena a nota obtida das avaliações objetivas realizadas na plataforma digital *on-line* e, por último, o atributo “Avaliações”, utilizado em 5 artigos, armazena a nota obtida na avaliação escrita.

Isidro et al. (2018), Waheed et al. (2020), Tomasevic et al. (2020), Islam et al. (2020), Kang & Wang (2018), Kostopoulos et al. (2019b), Kostopoulos et al. (2018a), Kostopoulos et al. (2019a), Karlos et al. (2020) e Waheed et al. (2023) utilizam dados de desempenho, como notas obtidas, para a análise, sem especificar quais atividades são desenvolvidas. Consequentemente, não é possível identificar os atributos utilizados.

Q6. Qual é o nível de escolaridade do público alvo nos estudos de evasão em EaD?

A preocupação com a evasão de estudantes atinge vários níveis de formação, do ensino médio ao superior. Por isso, as instituições de ensino adotaram estratégias de combate que permeiam entre a predição, detecção, diagnóstico ou acompanhamento. Os estudos dos 55 artigos primários trouxeram números relevantes sobre essa questão.

De acordo com a tabela 16, que apresenta o nível de escolaridade, quantidade de artigos e referências, 32 artigos estudam a evasão em EaD na graduação, pois os estudantes aceitam mais facilmente as plataformas de ensino como ferramentas de aprendizagem e compreendem que o mínimo de conhecimento de tecnologia e um equipamento com acesso a internet é suficiente para a participação nessa modalidade de ensino. No ensino médio técnico, 4 artigos investigam a evasão em EaD. Borrella et al. (2019) e Shafiq et al. (2022) pesquisam na pós-graduação e 18 artigos não informam onde realizaram a pesquisa, no entanto, agem na tentativa de mitigar a evasão em EaD. Logo, a pesquisa mostra que o alto índice de evasão do ensino médio à pós-graduação nessa modalidade de ensino é motivo de preocupação das instituições.

Tabela 16: Relação de artigos por nível de escolaridade do público alvo.

Escolaridade	Quantidade	Referências
Ensino Médio Técnico	4	Silva et al. (2015); Queiroga et al. (2017); Queiroga et al. (2022); Tamada et al. (2022)
Ensino Superior – Graduação	32	Kostopoulos et al. (2018b); Brandão et al. (2019); Ortigosa et al. (2019); Almeida Neto & Castro (2015); La Peña et al. (2017); Macedo et al. (2019); Isidro et al. (2018); Heidrich et al. (2018); Chen & Zhang (2017); Kang & Wang (2018); Kostopoulos et al. (2015); Kostopoulos et al. (2019b); Kostopoulos et al. (2018a); Kostopoulos et al. (2019a); Santos et al. (2015); Santos et al. (2016); Ramos et al. (2017); Rabelo et al. (2017); Ramos et al. (2018); Goel & Goyal (2020); Karlos et al. (2020); Lemay & Doleck (2020); Oreshin et al. (2020); Shafiq et al. (2022); Leite et al. (2021); Esteban et al. (2021); Nascimento et al. (2021); Adnan et al. (2022); Feldman-Maggor et al. (2022); Godinez & Lomibao (2022); Li et al. (2022); Waheed et al. (2023)
Ensino Superior – Pós-graduação	2	Borrella et al. (2019); Shafiq et al. (2022)
Não especificado	18	Brito et al. (2019a); Liang et al. (2016); Mishra & Mishra (2018); Cobos & Olmos (2019); Wang & Wang (2019); Oeda & Hashimoto (2017); Waheed et al. (2020); Tomasevic et al. (2020); Islam et al. (2020); Whitehill et al. (2017); Wang et al. (2017); Imran et al. (2019); Niu et al. (2018); Wu et al. (2019); Santos & Falcão (2017); Queiroga et al. (2019); Ng & Lei (2022); Tran et al. (2023)

Q7. Metodologias Ativas foram utilizadas para mitigar a evasão em EaD?

Com a análise dos 55 artigos primários estudados, podemos constatar que em nenhum dos trabalhos foram utilizadas Metodologias Ativas na abordagem de evasão e melhoria na permanência em EaD após a predição, detecção, diagnóstico ou acompanhamento do estudante do ensino médio à pós-graduação com técnicas, algoritmos e aplicações de MDE e LA em ambientes virtuais de aprendizagem.

Alguns artigos relatam o uso da metodologia para auxiliar professores no processo de ensino e aprendizagem e, conseqüentemente, no combate a evasão em EaD, contudo em situações em que a identificação não ocorreu por MDE e LA. Vale ressaltar que esses artigos não foram identificados conforme *strings* de busca.

O trabalho de Revathy & Kamalakkannan (2021) merece destaque, embora não esteja entre os artigos primários estudados, já que não se trata de um estudo focado em EaD. Tal estudo é o que mais se aproxima da questão desta pesquisa, pois apresenta uma proposta de *Framework* que instiga a aprendizagem colaborativa para alunos identificados como propensos a evasão através de Aprendizado de Máquina, no entanto, os autores não citam qual estratégia de aprendizagem e nem quais técnicas são utilizadas. Na mesma linha, Kostopoulos et al. (2019b) descrevem, como um aspecto interessante, a implementação de técnicas de aprendizado semi-supervisionado e aprendizado ativo no campo educacional, para prever o desempenho e as taxas de evasão dos alunos em instituições de ensino. A eficácia e a dinâmica dessas abordagens induzem a modelos preditivos ainda mais precisos e poderosos para a descoberta de conhecimento em dados educacionais.

Por conseguinte, a utilização de Metodologias Ativas favorece um relacionamento mais efetivo entre os atores e colabora para a diminuição do índice de evasão dos cursos oferecidos em EaD. Também, ajuda os professores nas práticas de ensino e os tornam mais proativos, capazes de acompanhar o desempenho dos alunos com medidas e ações preventivas em busca de resultados positivos e mitigar os riscos de desistência. Segundo Rigo et al. (2014), a análise do conjunto de informações existentes fornece pistas para buscar metodologias, para minimizar a evasão e potencializar a permanência dos estudantes em EaD.

6 Conclusão

Este trabalho apresentou um Mapeamento Sistemático da Literatura sobre o cenário mundial de pesquisas que envolvem o uso de Metodologias Ativas para mitigar a evasão e motivar a permanência de alunos identificados através de técnicas de Mineração de Dados Educacionais e *Learning Analytics* em cursos EaD. O mapeamento sistemático de 55 artigos demonstrou o crescente interesse pela questão da evasão nessa modalidade de ensino.

O alto índice de evasão de alunos em cursos oferecidos na modalidade a distância preocupa os gestores e professores das instituições de ensino, que buscam alternativas para identificar situações que motivam os alunos a permanecerem em seus cursos. Dentre as opções, o uso de Metodologias Ativas tem sido inserido em ambientes virtuais de ensino, pois apresentam fatores de melhoria no aprendizado do aluno e promovem interação, comunicação, desenvolvimento do senso crítico e autoaprendizagem.

Por meio da análise realizada, constatou-se que alguns trabalhos utilizam Metodologias Ativas para auxiliar os professores no processo de ensino e aprendizagem, na tentativa de diminuir o índice de evasão dos cursos e potencializar sua permanência. No entanto, nenhum dos estudos considerou a integração das Metodologias Ativas com as técnicas de MDE e LA, de forma a reduzir os riscos de abandono. Ainda sim, alguns trabalhos consideram a possibilidade de agregá-las às técnicas de mineração de dados e análise de aprendizagem, para contribuir com

a mitigação da evasão e aumento da permanência de alunos (Chandrasekaran et al., 2016; Lima & Siebra, 2017).

A utilização de MDE e LA para a análise de dados educacionais pode subsidiar a tomada de decisão, colaborar para identificar os riscos de evasão e potencializar a permanência dos alunos da EaD. Além disso, auxiliar o professor a acompanhar o desempenho do aluno durante o processo de aprendizagem, fator que pode amenizar, também, o problema da evasão.

Com o mapeamento foi possível identificar: (i) 5 técnicas, com ênfase em classificação e predição; (ii) 38 ferramentas amplamente utilizadas e 11 novas propostas desenvolvidas; (iii) 79 algoritmos utilizados e 13 novos desenvolvidos; (iv) os atributos que armazenam registros de dados demográficos, comportamento, interação e desempenho são frequentemente utilizados, com destaque para os dados de interação presentes em mais de 76% dos estudos; (v) mais de 58% dos artigos focaram a pesquisa no ensino superior, nível de graduação.

Cumprir destacar que, ainda, há um longo caminho a percorrer no sentido de diminuir o índice de evasão em EaD, seja com a ampliação das formas de controle, ou pela implementação de tecnologias, ou por ações de intervenção humana, ou pela aplicação e desenvolvimento de novas metodologias que favoreçam o aprendizado, aumentem o engajamento e a motivação dos alunos dessa modalidade de ensino.

Como trabalhos futuros pretende-se: (i) identificar as Metodologias Ativas que possam ser aplicadas no contexto da mitigação da evasão e promover a permanência em EaD; (ii) desenvolver um Sistema de Recomendação que utiliza Metodologias Ativas para melhorar o desempenho e mitigar os riscos de reprovação e evasão dos alunos em EaD.

Referências

- Acosta, O. C.; Reategui, E. B. & Behar, P. A. (2018). Recomendação de conteúdo em um ambiente colaborativo de Aprendizagem Baseada em Projetos. *Revista Brasileira de Informática na Educação*, 26(1), 91-111. doi: [10.5753/rbie.2018.26.01.91](https://doi.org/10.5753/rbie.2018.26.01.91) [GS Search]
- Adnan, M.; Uddin, M. I.; Khan, E.; Alharithi, F. S.; Amin, S. & Alzahrani, A. A. (2022). Earliest Possible Global and Local Interpretation of Students Performance in Virtual Learning Environment by Leveraging Explainable AI. *IEEE Access*, 10, 129843-129864. doi: [10.1109/ACCESS.2022.3227072](https://doi.org/10.1109/ACCESS.2022.3227072) [GS Search]
- Aldowah, H.; Al-Samarraie, H. & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13-49. doi: [10.1016/j.tele.2019.01.007](https://doi.org/10.1016/j.tele.2019.01.007) [GS Search]
- Almeida, C. M. M.; Scheunemann, C. M. B.; Santos, M. J. & Lopes, P. T. C. (2019). Propostas de metodologias ativas utilizando Tecnologias Digitais e ferramentas metacognitivas para auxiliar no processo de ensino e aprendizagem. *Revista Paradigma*, 40, 204-220. [GS Search]
- Almeida Neto, F. A. & Castro, A. (2015). Elicited and mined rules for dropout prevention in online courses. *IEEE Frontiers in Education Conference (FIE)*. IEEE. doi: [10.1109/FIE.2015.7344048](https://doi.org/10.1109/FIE.2015.7344048) [GS Search]
- Alves, M. O.; Medeiros, F. P. A.; Melo, L. B.; Barbosa, A. S. R. & Brito, M. L. Q. (2020b). Systematic Literature Review on the adoption of the Problem Based Learning methodology in Distance Education. *Iberian Conference on Information Systems and Technologies (CISTI)*, 1-4. IEEE. doi: [10.23919/CISTI49556.2020.9141089](https://doi.org/10.23919/CISTI49556.2020.9141089) [GS Search]

- Araújo, I. S. & Mazur, E. (2013). Instrução pelos colegas e ensino sob medida: uma proposta para o engajamento dos alunos no processo de ensino-aprendizagem de física. *Caderno Brasileiro de Ensino de Física*, 30(2), 362-384. doi: [10.5007/2175-7941.2013v30n2p362](https://doi.org/10.5007/2175-7941.2013v30n2p362) [GS Search]
- Baker, R. & Yacef, K. (2009). The state of educational data mining in 2009: a review and future visions. *Journal of Educational Data Mining (JEDM)*, 1(1), 3-17. doi: [10.5281/zenodo.3554657](https://doi.org/10.5281/zenodo.3554657) [GS Search]
- Baker, R.; Isotani, S. & Carvalho, A. (2011). Mineração de Dados Educacionais: Oportunidades para o Brasil. *Revista Brasileira de Informática na Educação*, 19(2), 3-13. doi: [10.5753/rbie.2011.19.02.03](https://doi.org/10.5753/rbie.2011.19.02.03) [GS Search]
- Bergmann, J. & Sams, A. (2016). Sala de Aula Invertida: uma metodologia ativa de aprendizagem. LTC. [GS Search]
- Borrella, I.; Caballero-Caballero, S. & Ponce-Cueto, E. (2019). Predict and Intervene: Addressing the Dropout Problem in a MOOC-based Program. *ACM Conference on Learning*, 1-9. ACM. doi: [10.1145/3330430.3333634](https://doi.org/10.1145/3330430.3333634) [GS Search]
- Brandão, I. V.; da Costa, J. P. C. L.; Santos, G. A.; Praciano, B. J. G.; D. Junior, F. C. M. & S. Junior, R. T. (2019). Classification and predictive analysis of educational data to improve the quality of distance learning courses. *Workshop on Communication Networks and Power Systems (WCNPS)*. IEEE. doi: [10.1109/WCNPS.2019.8896312](https://doi.org/10.1109/WCNPS.2019.8896312) [GS Search]
- Brito, M.; Medeiros, F. & Bezerra, E. P. (2019a). An Infographics-based Tool for Monitoring Dropout Risk on Distance Learning in Higher Education. *International Conference on Information Technology Based Higher Education and Training (ITHET)*. IEEE. doi: [10.1109/ITHET46829.2019.8937361](https://doi.org/10.1109/ITHET46829.2019.8937361) [GS Search]
- Brito, M. T. S. (2019b). Um plugin do tipo report para a identificação do risco de evasão na educação superior a distância que usa técnicas de visualização de dados. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Informática, Universidade Federal da Paraíba, João Pessoa, 2019. [GS Search]
- Cambruzzi, W.; Rigo, S. J. & Barbosa, J. L. V. (2015). Dropout Prediction and Reduction in Distance Education Courses with the Learning Analytics Multitrail Approach. *Journal of Universal Computer Science*, 21(1), 23-47. [GS Search]
- Censo EAD.BR: relatório analítico da aprendizagem a distância no Brasil 2020. Censo EAD.BR: Analytic Report of Distance Learning in Brazil 2020 [livro eletrônico]/[organização ABED – Associação Brasileira de Educação a Distância. Camila Rosa (tradutora). Curitiba: InterSaberes, 2022. [GS Search]
- Chandrasekaran, D.; Thirunavukkarasu, G. S. & Littlefair, G. (2016). Collaborative Learning Experience of Students in Distance Education. *International Symposium on Project Approaches in Engineering Education and Active Learning in Engineering Education Workshop*, 90-99. [GS Search]
- Chen, Y. & Zhang, M. (2017). MOOC Student Dropout: Pattern and Prevention. *ACM Turing 50th Celebration Conference*, 1-6. ACM. doi: [10.1145/3063955.3063959](https://doi.org/10.1145/3063955.3063959) [GS Search]
- Cobos, R. & Olmos, L. (2019). A Learning Analytics Tool for Predictive Modeling of Dropout and Certificate Acquisition on MOOCs for Professional Learning. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 1533-1537. IEEE. doi: [10.1109/IEEM.2018.8607541](https://doi.org/10.1109/IEEM.2018.8607541) [GS Search]

- Correia, C. F. & Pimentel, E. P. (2011). Mineração de dados na formação de turmas para a recuperação paralela na Educação Básica. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 172–175. doi: [10.5753/cbie.sbie.2011.%25p](https://doi.org/10.5753/cbie.sbie.2011.%25p) [GS Search]
- Esteban, A.; Romero, C. & Zafra, A. (2021). Assignments as Influential Factor to Improve the Prediction of Student Performance in Online Courses. *Applied Sciences*, 11(21), 10145. doi: [10.3390/app112110145](https://doi.org/10.3390/app112110145) [GS Search]
- Feldman-Maggor, Y.; Blonder, R. & Tuvi-Arad, I. (2022). Let them choose: Optional assignments and online learning patterns as predictors of success in online general chemistry courses. *The Internet and Higher Education*, 55, 100867. Elsevier. doi: [10.1016/j.iheduc.2022.100867](https://doi.org/10.1016/j.iheduc.2022.100867) [GS Search]
- Fonseca, S. M. & Mattar, J. (2017). Metodologias ativas aplicadas à educação a distância: revisão de literatura. *Revista EDAPECI: educação a distância e práticas educativas comunicacionais e interculturais*, 17(2), 185-197. doi: [10.29276/redapeci.2017.17.26509.185-197](https://doi.org/10.29276/redapeci.2017.17.26509.185-197) [GS Search]
- Godinez, C. & Lomibao, L. (2022). A Gaussian-Bernoulli Mixed Naïve Bayes Approach to Predict Students' Academic Procrastination Tendencies in Online Mathematics Learning. *American Journal of Educational Research*, 10(4), 223-232. doi: [10.12691/education-10-4-10](https://doi.org/10.12691/education-10-4-10) [GS Search]
- Goel, Y. & Goyal, R. (2020). On the Effectiveness of Self-Training in MOOC Dropout Prediction. *Open Computer Science*, 10, 246-258. doi: [10.1515/comp-2020-0153](https://doi.org/10.1515/comp-2020-0153) [GS Search]
- Heidrich, L.; Barbosa, J. L. Victória; Cambuzzi, W.; Rigo, S. J.; Martins, M. G. & Santos, R. B. S. (2018). Diagnosis of learner dropout based on learning styles for online distance learning. *Telematics and Informatics*, 35(6), 1593-1606. Elsevier. doi: [10.1016/j.tele.2018.04.007](https://doi.org/10.1016/j.tele.2018.04.007) [GS Search]
- Hoffmann, A. T.; Jacques, J. J.; Silva, T. L. K. & Silva, R. P. (2020). Revisão sistemática da literatura: metodologias ativas de ensino-aprendizagem e sua utilização nos cursos de design, engenharia e arquitetura. *Design em Pesquisa*, 34-54. Marca visual. [GS Search]
- Imran, A. S.; Dalipi, F. & Kastrati, Z. (2019). Predicting Student Dropout in a MOOC: An Evaluation of a Deep Neural Network Model. *International Conference on Computing and Artificial Intelligence*, 190-195. ACM. doi: [10.1145/3330482.3330514](https://doi.org/10.1145/3330482.3330514) [GS Search]
- Isidro, C.; Carro, R. M. & Ortigosa, A. (2018). Dropout Detection in MOOCs: An Exploratory Analysis. *International Symposium on Computers in Education (SIIE)*. IEEE. doi: [10.1109/SIIE.2018.8586748](https://doi.org/10.1109/SIIE.2018.8586748) [GS Search]
- Islam, O.; Siddiqui, M. & Aljohani, N. R. (2020). Identifying Online Profiles of Distance Learning Students Using Data Mining Techniques. *International Conference on Digital Technology in Education (ICDTE)*, 115-120. ACM. doi: [10.1145/3369199.3369249](https://doi.org/10.1145/3369199.3369249) [GS Search]
- Kang, K. & Wang, S. (2018). Analyze and Predict Student Dropout from Online Programs. *International Conference on Compute and Data Analysis*, 6-12. ACM. doi: [10.1145/3193077.3193090](https://doi.org/10.1145/3193077.3193090) [GS Search]
- Karlos, S.; Kostopoulos, G. & Kotsiantis, S. (2020). Predicting and Interpreting Students' Grades in Distance Higher Education through a Semi-Regression Method. *Applied Sciences*, 10(23), 8413. doi: [10.3390/app10238413](https://doi.org/10.3390/app10238413) [GS Search]

- Kostopoulos, G.; Karlos, S. & Kotsiantis, S. (2019a). Multiview Learning for Early Prognosis of Academic Performance: A Case Study. *IEEE Transactions on Learning Technologies*, 12(2), 212-224. doi: [10.1109/TLT.2019.2911581](https://doi.org/10.1109/TLT.2019.2911581) [GS Search]
- Kostopoulos, G.; Kotsiantis, S.; Fazakis, N.; Koutsonikos, G. & Pierrakeas, C. (2019b). A Semi-Supervised Regression Algorithm for Grade Prediction of Students in Distance Learning Courses. *International Journal on Artificial Intelligence Tools*, 28, 1-19. doi: [10.1142/S0218213019400013](https://doi.org/10.1142/S0218213019400013) [GS Search]
- Kostopoulos, G.; Kotsiantis, S.; Pierrakeas, C.; Koutsonikos, G. & Gravvanis, G.A. (2018a). Forecasting students' success in an open university. *International Journal Learning Technology*, 13, 26-43. doi: [10.1504/IJLT.2018.091630](https://doi.org/10.1504/IJLT.2018.091630) [GS Search]
- Kostopoulos, G.; Kotsiantis, S. & Pintelas, P. (2015). Estimating student dropout in distance higher education using semi-supervised techniques. *Panhellenic Conference on Informatics (PCI 2015)*, 38-43. doi: [10.1145/2801948.2802013](https://doi.org/10.1145/2801948.2802013) [GS Search]
- Kostopoulos, G.; Kotsiantis, S.; Ragos, O. & Grapsa, T.N. (2018b). Early dropout prediction in distance higher education using active learning. *International Conference on Information, Intelligence, Systems and Applications (IISA 2017)*, 1-6. doi: [10.1109/IISA.2017.8316424](https://doi.org/10.1109/IISA.2017.8316424) [GS Search]
- La Peña, D.; Lara, J. A.; Lizcano, D.; Martínez, M. A.; Burgos, C. & Campanario, M. L. (2017). Mining activity grades to model students' performance. *International Conference on Engineering & MIS (ICEMIS)*. IEEE. doi: [10.1109/ICEMIS.2017.8272963](https://doi.org/10.1109/ICEMIS.2017.8272963) [GS Search]
- Leite, D.; Filho, E.; Oliveira, J. F. L.; Carneiro, R. E. & Maciel, A. (2021). Early detection of students at risk of failure from a small dataset. *International Conference on Advanced Learning Technologies (ICALT)*, 42-46. IEEE. doi: [10.1109/ICALT52272.2021.00021](https://doi.org/10.1109/ICALT52272.2021.00021) [GS Search]
- Lemay, D. & Doleck, T. (2020). Predicting completion of massive open online course (MOOC) assignments from video viewing behavior. *Interactive Learning Environments*, 30, 1782-1793. doi: [10.1080/10494820.2020.1746673](https://doi.org/10.1080/10494820.2020.1746673) [GS Search]
- Li, Y.; Cui, X. & Zhang, Z. (2022). Dropout Rate Prediction for MOOC based on Inceptiontime Model. *International Conference on Distance Education and Learning (ICDEL)*, 54-59. doi: [10.1145/3543321.3543330](https://doi.org/10.1145/3543321.3543330) [GS Search]
- Liang, J.; Yang, J.; Wu, Y.; Li, C. & Zheng, L. (2016). Big Data Application in Education: Dropout Prediction in Edx MOOCs. *IEEE Second International Conference on Multimedia Big Data (BigMM)*, 440-443. IEEE. doi: [10.1109/BigMM.2016.70](https://doi.org/10.1109/BigMM.2016.70) [GS Search]
- Lima, E. & Siebra, C. (2017). CollabEduc: Uma Ferramenta de Colaboração em Pequenos Grupos para Plataformas de Aprendizagem a Distância. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1707-1716. doi: [10.5753/cbie.sbie.2017.1707](https://doi.org/10.5753/cbie.sbie.2017.1707) [GS Search]
- Macedo, M.; Santana Jr, C.; Siqueira, H.; Rodrigues, R. L., Ramos, J. L. C.; Silva, J. C. S.; Maciel, A. M. A. & Bastos-Filho, C. J. A. (2019). Investigation of College Dropout with the Fuzzy C-Means Algorithm. *International Conference on Advanced Learning Technologies (ICALT)*, 187-189. IEEE. doi: [10.1109/ICALT.2019.00055](https://doi.org/10.1109/ICALT.2019.00055) [GS Search]
- Manhães, L. M. B.; Da Cruz, S. M. S.; Costa, R. J. M.; Zavaleta, J. & Zimbrão, G. (2011). Previsão de estudantes com risco de evasão utilizando técnicas de mineração de dados. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 150-159. doi: [10.5753/cbie.sbie.2011.%25p](https://doi.org/10.5753/cbie.sbie.2011.%25p) [GS Search]

- Mattar, J. (2017). Metodologias ativas para a educação presencial, blended e a distância. *Artesanato Educacional*. [[GS Search](#)]
- Mishra, B. B. & Mishra, S. (2018). Quality Improvements in Online Education System by Using Data Mining Techniques. *International Conference on Data Science and Business Analytics (ICDSBA)*, 532-536. IEEE. doi: [10.1109/ICDSBA.2018.00105](#) [[GS Search](#)]
- Mota, A. R. & Rosa, C. T. W. (2018). Ensaio sobre metodologias ativas: reflexões e propostas. *Revista Espaço Pedagógico*, 25(2), 261-276. doi: [10.5335/rep.v25i2.8161](#) [[GS Search](#)]
- Nascimento, P. S. C.; Silva Junior, A. S.; Schulz, C. L.; Santos, M. V. R.; Maciel, A. M. A.; Rodrigues, R. L.; Nascimento, R. R. & Alencar, F. M. R. (2021). Análise dos Impactos da Gestão do Tempo no Desempenho Acadêmico Através da Mineração de Dados Educacionais. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 783-791. doi: [10.5753/sbie.2021.217742](#) [[GS Search](#)]
- Ng, K. & Lei, P. (2022). A Lightweight Method using LightGBM Model with Optuna inMOOCs Dropout Prediction. *International Conference on Education and Multimedia Technology (ICMT)*, 53-59. ACM. doi: [10.1145/3551708.3551732](#) [[GS Search](#)]
- Niu, Z.; Li, W.; Yan, X. & Wu, N. (2018). Exploring Causes for the Dropout on Massive Open Online Courses. *ACM Turing Celebration Conference*, 47-52. ACM. doi: [10.1145/3210713.3210727](#) [[GS Search](#)]
- Oeda, S. & Hashimoto, G. (2017). Log-Data Clustering Analysis for Dropout Prediction in Beginner Programming Classes. *Procedia Computer Science*, 112, 614-621. ACM. doi: [10.1016/j.procs.2017.08.088](#) [[GS Search](#)]
- Oreshin, S.; Filchenkov, A.; Petrusha, P.; Krashennnikov, E.; Panfilov, A.; Glukhov, I.; Kaliberda, Y.; Masalskiy, D.; Serdyukov, A.; Kazakovtsev, V.; Khlopotov, M.; Podolenchuk, T.; Smetannikov, I. & Kozlova, D. (2020). Implementing a Machine Learning Approach to Predicting Students' Academic Outcomes. *International Conference on Control, Robotics and Intelligent System (CCRIS)*, 78-83. doi: [10.1145/3437802.3437816](#) [[GS Search](#)]
- Ortigosa, A.; Carro, R. M.; Bravo-Agapito J.; Lizcano, D.; Alcolea, J. J. & Blanco, O. (2019). From Lab to Production: Lessons Learnt and Real-Life Challenges of an Early Student-Dropout Prevention System. *IEEE Transactions on Learning Technologies*, 12(2), 264-277. IEEE. doi: [10.1109/TLT.2019.2911608](#) [[GS Search](#)]
- Petersen, K.; Vakkalanka, S. & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64, 1-18. Elsevier. doi: [10.1016/j.infsof.2015.03.007](#) [[GS Search](#)]
- Queiroga, E.; Cechinel, C. & Araújo, R. (2017). Predição de estudantes com risco de evasão em cursos técnicos à distância. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1547-1556. doi: [10.5753/cbie.sbie.2017.1547](#) [[GS Search](#)]
- Queiroga, E. M.; Cechinel, C. & Aguiar, M. S. (2019). Uma abordagem para predição de estudantes em risco utilizando algoritmos genéticos e mineração de dados: um estudo de caso com dados de um curso técnico a distância. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 119-128. doi: [10.5753/cbie.wcbie.2019.119](#) [[GS Search](#)]
- Queiroga, E. M.; Paragarino, V. R.; Casas, A. P.; Primo, T. T.; Munoz, R.; Ramos, V. C. & Cechinel, C. (2022). Experimenting Learning Analytics and Educational Data Mining in different educational contexts and levels. *Latin American Conference on Learning Technologies (LACLO)*, 1-9. IEEE. doi: [10.1109/LACLO56648.2022.10013478](#) [[GS Search](#)]

- Rabelo, H.; Burlamaqui, A. M. F.; Valentim, R. A. M.; Rabelo, D. S. S. & Medeiros, S. R. S. (2017). Utilização de técnicas de Mineração de Dados Educacionais para a predição de desempenho de alunos de EaD em Ambientes Virtuais de Aprendizagem. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1527-1536. doi: [10.5753/cbie.sbie.2017.1527](https://doi.org/10.5753/cbie.sbie.2017.1527) [GS Search]
- Ramos, J. L. C.; Silva, J. C. S.; Prado, L. C.; Gomes, A. S.; Souza, F. F. D.; Zambom, E. G. & Rodrigues, R. L. (2017). Um Modelo Preditivo da Evasão dos Alunos na EAD a partir dos Construtos da Teoria da Distância Transacional. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1227-1236. doi: [10.5753/cbie.sbie.2017.1227](https://doi.org/10.5753/cbie.sbie.2017.1227) [GS Search]
- Ramos, J. L. C.; Silva, J. C. S.; Prado, L. C.; Gomes, A. S. & Rodrigues, R. L. (2018). Um estudo comparativo de classificadores na previsão da evasão de alunos em EAD. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1463-1472. doi: [10.5753/cbie.sbie.2018.1463](https://doi.org/10.5753/cbie.sbie.2018.1463) [GS Search]
- Revathy, M. & Kamalakkannan, S. (2021). Collaborative learning for improving intellectual skills of dropout students using datamining techniques. *International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 236-240. IEEE. doi: [10.1109/ICAIS50930.2021.9395912](https://doi.org/10.1109/ICAIS50930.2021.9395912) [GS Search]
- Rigo, S. J.; Cambruzzi, W.; Barbosa, J. L. V. & Cazella, S. C. (2014). Aplicações de Mineração de Dados Educacionais e Learning Analytics com foco na evasão escolar: oportunidades e desafios. *Revista Brasileira de Informática na Educação*, 22(1), 132-146. doi: [10.5753/rbie.2014.22.01.132](https://doi.org/10.5753/rbie.2014.22.01.132) [GS Search]
- Rodrigues, M.W.; Isotani, S. & Zárata, L. E. (2018). Educational Data Mining: A review of evaluation process in the e-learning. *Telematics and Informatics*, 35(6), 1701–1717. doi: [10.1016/j.tele.2018.04.015](https://doi.org/10.1016/j.tele.2018.04.015) [GS Search]
- Santos, F. D.; Bercht, M. & Wives, L. K. (2015). Classificação de alunos desanimados em um AVEA: uma proposta a partir da mineração de dados educacionais. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1052- 1061. doi: [10.5753/cbie.sbie.2015.1052](https://doi.org/10.5753/cbie.sbie.2015.1052) [GS Search]
- Santos, R. M. M.; Pitangui, C. G.; Andrade, A. V. & Assis, L. P. (2016). Uso de Séries Temporais e Seleção de Atributos em Mineração de Dados Educacionais para Previsão de Desempenho Acadêmico. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1146-1155. doi: [10.5753/cbie.sbie.2016.1146](https://doi.org/10.5753/cbie.sbie.2016.1146) [GS Search]
- Santos, D. C. V. & Falcão, T. P. (2017). Acompanhamento de Alunos em Ambientes Virtuais de Aprendizagem Baseado em Sistemas Tutores Inteligentes. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1267-1276. doi: [10.5753/cbie.wcbie.2017.90](https://doi.org/10.5753/cbie.wcbie.2017.90) [GS Search]
- Santos, J. F.; Sousa, J. D. A.; Mello, R. F.; Cristino, C. T. & Alves, G. (2021a). Um modelo para análise do impacto de retenção e evasão no ensino superior utilizando Cadeias de Markov Absorventes. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 813-823. doi: [10.5753/sbie.2021.218219](https://doi.org/10.5753/sbie.2021.218219) [GS Search]
- Schlemmer, E. (2014). Gamificação em espaços de convivência híbridos e multimodais: design e cognição em discussão. *Revista FAEBA*, 23(42), 73-89. doi: [10.21879/faeaba2358-0194.2014.v23.n42.p%p](https://doi.org/10.21879/faeaba2358-0194.2014.v23.n42.p%p) [GS Search]
- Shafiq, D. A.; Marjani, M.; Habbeb, R. A. A. & Asirvatham, D. (2022). A Conceptual Predictive Analytics Model for the Identification of at-risk students in VLE using Machine Learning Techniques. *International Conference on Mathematics, Actuarial Science*,

- Computer Science and Statistics (MACS)*, 1-8. IEEE. doi: [10.1109/MACS56771.2022.10023143](https://doi.org/10.1109/MACS56771.2022.10023143) [GS Search]
- Siemens, G. & d Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. *International Conference on Learning Analytics and Knowledge*, 252–254. ACM. doi: [10.1145/2330601.2330661](https://doi.org/10.1145/2330601.2330661) [GS Search]
- Silva, F.; Da Silva, J.; Silva, R. & Fonseca, L. C. (2015). Um modelo preditivo para diagnóstico de evasão baseado nas interações de alunos em fóruns de discussão. *Simpósio Brasileiro de Informática na Educação (SBIE)*, 1187-1196. doi: [10.5753/cbie.sbie.2015.1187](https://doi.org/10.5753/cbie.sbie.2015.1187) [GS Search]
- Silva, J. C. & Castro, M. C. D. (2022). The challenges of academic retention in Distance Learning: identification and analysis of Critical Success Factors. *Research, Society and Development*, 11(12), 1-16. doi: [10.33448/rsd-v11i12.34078](https://doi.org/10.33448/rsd-v11i12.34078) [GS Search]
- Tamada, M.; Giusti, R. & Netto, J. (2022). Predicting Students at Risk of Dropout in Technical Course Using LMS Logs. *Electronics*, 11(3), 468. doi: [10.3390/electronics11030468](https://doi.org/10.3390/electronics11030468) [GS Search]
- Tomasevic, N.; Gvozdenovic, N. & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education*, 143, 1-18. Elsevier. doi: [10.1016/j.compedu.2019.103676](https://doi.org/10.1016/j.compedu.2019.103676) [GS Search]
- Tran, T. P.; Jan, T. & Kew, S. N. (2023). Learning Analytics for Improved Course Delivery: Applications and Techniques. *International Conference on Digital Technology in Education (ICDTE)*, 100-106. doi: [10.1145/3568739.3568758](https://doi.org/10.1145/3568739.3568758) [GS Search]
- Waheed, H.; Hassan, S.; Aljohani, N. R.; Hardman, J.; Alelyani, S. & Nawaz, R. (2020). Predicting academic performance of students from VLE big data using deep learning models. *Computers in Human Behavior*, 104(1), 1-13. Elsevier. doi: [10.1016/j.chb.2019.106189](https://doi.org/10.1016/j.chb.2019.106189) [GS Search]
- Waheed, H.; Hassan, S.; Nawaz, R.; Aljohani, N. R.; Chen, G. & Gasevic, D. (2023). Early prediction of learners at risk in self-paced education: A neural network approach. *Expert Systems with Applications*, 213(A), 118868. Elsevier. doi: [10.1016/j.eswa.2022.118868](https://doi.org/10.1016/j.eswa.2022.118868) [GS Search]
- Wang, W.; Yu, H. & Miao, C. (2017). Deep Model for Dropout Prediction in MOOCs. *International Conference on Crowd Science and Engineering*, 26-32. ACM. doi: [10.1145/3126973.3126990](https://doi.org/10.1145/3126973.3126990) [GS Search]
- Wang, L. & Wang, H. (2019). Learning Behavior Analysis and Dropout Rate Prediction Based on MOOCs Data. *International Conference on Information Technology in Medicine and Education (ITME)*, 419-423. IEEE. doi: [10.1109/ITME.2019.00100](https://doi.org/10.1109/ITME.2019.00100) [GS Search]
- Widyahastuti, F. & Tjhin, U. (2018). Performance Prediction in Online Discussion Forum: state-of-the-art and comparative analysis. *International Conference on Computer Science and Computational Intelligence*, 302-314. Elsevier. doi: [10.1016/j.procs.2018.08.178](https://doi.org/10.1016/j.procs.2018.08.178) [GS Search]
- Whitehill, J.; Mohan, K.; Seaton, D.; Rosen, Y. & Tingley, D. (2017). MOOC Dropout Prediction: How to Measure Accuracy? *Fourth (2017) ACM Conference on Learning*, 161-164. ACM. doi: [10.1145/3051457.3053974](https://doi.org/10.1145/3051457.3053974) [GS Search]
- Wu, N.; Zhang, L.; Gao, Y.; Zhang, M.; Sun, X. & Feng, J. (2019). CLMS-Net: Dropout Prediction in MOOCs with Deep Learning. *ACM Turing Celebration Conference*, 1-6. ACM. doi: [10.1145/3321408.3322848](https://doi.org/10.1145/3321408.3322848) [GS Search]