

Mineração de Dados Educacionais e Análise de Sentimentos em Ambientes Virtuais de Aprendizagem: um Mapeamento Sistemático

Educational Data Mining and Sentiment Analysis in Virtual Learning Environments: a Systematic Mapping

ISSN 2177-8310
DOI: 10.18264/eadf.v12i2.1786

Resumo

A análise de sentimento é uma área de Mineração de Dados que envolve processamento de linguagem natural, extração de informações, inteligência artificial e aprendizado de máquina. Assim, por meio dessa técnica e, também, das emoções dos alunos em ambientes virtuais de aprendizagem, é possível descobrir padrões e verificar eventuais deficiências durante o processo de aprendizagem. O objetivo deste artigo é apresentar os resultados de um mapeamento sistemático da literatura realizado sobre técnicas, métodos, algoritmos, bibliotecas e ferramentas de mineração de dados educacionais utilizados para análise de sentimentos e emoções dos estudantes em ambientes virtuais de aprendizagem. Além disso, foram identificadas as finalidades para a análise de sentimentos e os tipos de emoções consideradas. Portanto, foram selecionados 20 estudos primários para verificação em profundidade. Os resultados apresentam a predominância de algoritmos de aprendizado de máquina para análise de sentimentos, abordando avaliação de cursos e professores, a eficácia do ambiente de aprendizagem, a satisfação e as dificuldades dos alunos. Além disso, a maioria dos estudos explora a polaridade de sentimento: positivo, negativo e neutro.

Palavras-chave: Mineração de dados educacionais. Sentimentos. Emoções. Ambiente virtual de aprendizagem.

Rafael Leonardo Vivian^{1,2*}

Silvio César Cazella³

Leticia Rocha Machado¹

Patricia Alejandra Behar¹

¹ Universidade Federal do Rio Grande do Sul. Av. Paulo Gama, 110 – Porto Alegre – RS – Brasil.

² Instituto Federal Catarinense. Rua Cruz e Souza, 89 – Fraiburgo – SC – Brasil.

³ Universidade Federal de Ciências da Saúde de Porto Alegre. Rua Sarmiento Leite, 2457 – Porto Alegre – RS – Brasil.

*rafael.vivian@ifc.edu.br



Recebido 06/04/2022
Aceito 14/06/2022
Publicado 22/06/2022

COMO CITAR ESTE ARTIGO

ABNT: VIVIAN, R. L. et al. Mineração de Dados Educacionais e Análise de Sentimentos em Ambientes Virtuais de Aprendizagem: um Mapeamento Sistemático. **EAD em Foco**, v. 12, n. 2, e1786, 2022. doi: <https://doi.org/10.18264/eadf.v12i2.1786>

Educational Data Mining and Sentiment Analysis in Virtual Learning Environments: a Systematic Mapping

Abstract

Sentiment analysis is a Data Mining area that involves natural language processing, information extraction, artificial intelligence and machine learning. Thus, sentiment analysis and also emotions from students in virtual learning environments enable verifying possible deficiencies during the learning process, for example. The aim of this paper is to present the results of a systematic mapping of the literature carried out on techniques, methods, algorithms, libraries and tools about educational data mining used for analyzing sentiments and emotions from students in virtual learning environments. Furthermore, the purposes for the analysis of sentiments and the types of emotions considered were identified. Therefore, 20 primary studies were selected to verify details. The results show the predominance of machine learning algorithms for sentiment analysis, addressing courses and teachers evaluation, the effectiveness of the learning environment, student satisfaction and difficulties. Furthermore, most studies explore the sentiment polarity: positive, negative and neutral.

Keywords: Educational data mining. Sentiments. Emotions. Virtual learning environment.

1. Introdução

O aumento do uso de tecnologias da informação e comunicação em sistemas educacionais tem propiciado o armazenamento de grandes quantidades de dados de alunos, o que oportuniza o uso de Mineração de Dados Educacionais (MDE) para melhorar os processos de ensino e aprendizagem (HEUBNER, 2013; SACHIN e VIJAY, 2012). A MDE é útil em diversas áreas, incluindo a identificação de sentimentos e emoções, descobrindo as necessidades de aprendizagem para diferentes grupos de alunos (ARUNA, SASANKA e VINAY, 2021). Dessa forma, a análise de sentimento pode revelar uma relação entre o rendimento referente ao aprendizado e o interesse do aluno em um determinado curso.

A análise de sentimento é entendida como uma técnica de mineração de opinião e pode envolver: Processamento de Linguagem Natural (PLN), extração de informações, Inteligência Artificial, Aprendizado de Máquina e Mineração de Dados (PURAIVAN *et al.*, 2021). Assim, essa técnica é dividida em quatro fases: aquisição de dados, preparação de dados, análise de revisão e classificação de sentimento. Além disso, existem duas abordagens principais: (1) aprendizado de máquina, que é dividido em supervisionado e não supervisionado; e (2) abordagem baseada em léxico, que é dividida em duas categorias, baseadas em dicionário e baseadas em corpus (ANITHA, ANITHA e PRADEEPA, 2013).

Em Ambientes Virtuais de Aprendizagem (AVA), as interações entre alunos, professor e tutor ocorrem, principalmente, por meio de mensagens em fóruns e chats. As diversas disciplinas e os diferentes conteúdos dos cursos podem gerar uma quantidade de informação no AVA que torna trabalhoso para professores e tutores examinarem os textos manualmente para verificar os sentimentos dos alunos e para descobrir eventuais obstáculos durante o processo de ensino e aprendizagem.

Nesse cenário, o objetivo deste artigo é apresentar os resultados de um mapeamento sistemático da literatura sobre MDE e Análise de Sentimentos em AVA. Assim, este mapeamento sistemático permitiu identificar os principais aspectos da literatura sobre o assunto no período de 2010 a 2021, publicados em periódicos e eventos científicos, além de verificar oportunidades para trabalhos futuros.

O presente artigo está organizado da seguinte forma. A Seção 2 trata do procedimento de mapeamento sistemático aplicado; a Seção 3 descreve os resultados obtidos; a Seção 4 apresenta a discussão e, finalmente, a Seção 5 encerra com as conclusões e limitações.

2. Metodologia

O mapeamento sistemático da literatura é um processo de coleta, avaliação e sistematização de estudos primários. Assim, é apresentada uma abordagem mais ampla e quantitativa que fornece dados para identificação de tendências de estudos, bem como lacunas a serem exploradas (PETERSEN *et al.*, 2015). Nessa perspectiva, o foco deste mapeamento sistemático de literatura é identificar estudos primários sobre MDE e Análise de Sentimentos em AVA. Portanto, optou-se por este tipo de estudo para obter uma visão geral das áreas, no que diz respeito às técnicas, métodos, algoritmos, bibliotecas e ferramentas utilizadas, às finalidades da análise de sentimentos e às emoções e sentimentos analisados.

2.1. Questões de pesquisa

As questões de pesquisa (QP) foram elaboradas considerando o objetivo delineando, sendo apresentadas no Quadro 1, a fim de verificar as evidências na literatura científica sobre MDE, análise de sentimentos e AVA.

Quadro 1: Questões de pesquisa.

QP1	Quais são as técnicas, métodos, algoritmos, bibliotecas e ferramentas de mineração de dados educacionais para análise de base de dados referentes a sentimentos e emoções do estudante no processo de aprendizagem em AVA?
QP2	Com qual objetivo foi empregada a análise de sentimentos em AVA?
QP3	Quais tipos de emoções e sentimentos têm sido considerados nos estudos?

Fonte: elaborado pelos autores.

2.2. Processo de busca

A pesquisa foi realizada em bases de dados de literatura científica usando palavras-chave em inglês para as categorias “mineração de dados”, “análise de sentimentos” e “ambientes virtuais de aprendizagem”, conforme apresentadas no Quadro 2. Em cada categoria, foram combinadas as palavras-chave usando o operador booleano “OR”; ambas as categorias foram agrupadas usando o operador booleano “AND”. As seguintes bases de dados de literatura científica foram pesquisadas:

1. ACM Digital Library (<https://dl.acm.org>)
2. IEEE Xplore (<https://ieeexplore.ieee.org>)
3. Scopus (<https://www.scopus.com>)
4. Springer Link (<https://link.springer.com>)

Quadro 2: Palavras-chave usadas neste mapeamento.

Referência	Categoria	Palavras-chave
C1	Mineração de dados	<i>Data mining</i>
C2	Análise de sentimentos	<i>Sentiment analysis</i>
		<i>Emotion analysis</i>
C3	Ambiente Virtual de Aprendizagem	<i>Virtual learning environment</i>
		<i>E-learning</i>
		<i>Distance learning</i>
		<i>Distance education</i>
		<i>Online learning</i>
C4	Procedimentos	<i>Technique</i>
		<i>Method</i>
		<i>Algorithm</i>
		<i>Tool</i>

Fonte: elaborado pelos autores.

Assim, buscaram-se artigos científicos completos com foco em “mineração de dados”, “análise de sentimentos” e “ambientes virtuais de aprendizagem”, escritos em inglês¹, publicados entre Janeiro de 2010 e Setembro de 2021. A *string* de busca foi definida como uma combinação de C1, C2, C3 e C4, usando os operadores booleanos “AND” e “OR”, conforme apresentado no Quadro 3.

Quadro 3: *String* de busca aplicada em cada base de dados.

Base de dados	String
ACM Digital Library	<i>[All: “data mining”] AND [[All: “sentiment analysis”] OR [All: “emotion analysis”]] AND [[All: “virtual learning environment”] OR [All: “e-learning”] OR [All: “distance learning”] OR [All: “distance education”] OR [All: “online learning”]] AND [[All: “technique”] OR [All: “method”] OR [All: “algorithm”] OR [All: “tool”]]</i>
IEEE Xplore	<i>(“data mining”) AND (“sentiment analysis” OR “emotion analysis”) AND (“virtual learning environment” OR “e-learning” OR “distance learning” OR “distance education” OR “online learning”) AND (“technique” OR “method” OR “algorithm” OR “tool”)</i>
Scopus	<i>TITLE-ABS-KEY (“data mining”) AND (“sentiment analysis” OR “emotion analysis”) AND (“virtual learning environment” OR “e-learning” OR “distance learning” OR “distance education” OR “online learning”) AND (“technique” OR “method” OR “algorithm” OR “tool”)</i>
Springer Link	<i>(“data mining”) AND (“sentiment analysis” OR “emotion analysis”) AND (“virtual learning environment” OR “e-learning” OR “distance learning” OR “distance education” OR “online learning”) AND (“technique” OR “method” OR “algorithm” OR “tool”)</i>

Fonte: elaborado pelos autores.

A *string* de busca foi adaptada conforme a sintaxe de cada base de dados de literatura científica, uma vez que as opções de pesquisa diferem e a forma como deve ser construída é específica para cada base.

1 Optou-se pela língua inglesa pela possibilidade de acessar um volume maior de publicações científicas.

2.3. Critérios de inclusão e exclusão

Os critérios de inclusão e exclusão utilizados neste mapeamento sistemático aplicados na pesquisa são os listados no Quadro 4.

Quadro 4: Critérios de inclusão e exclusão.

Critérios de inclusão	Critérios de exclusão
CI1. Estudos primários CI2. Estudos que abordam Mineração de Dados Educacionais, Análise de Sentimentos e Ambientes Virtuais de Aprendizagem	CE1. Estudos secundários CE2. Livros, capítulos de livros, teses e dissertações de pós-graduação, manuais, relatórios CE3. Estudos duplicados (somente uma cópia de cada estudo foi incluído) CE4. Estudos a que não foi possível ter acesso CE5. Estudos não escritos na língua inglesa CE6. Estudos que não abordam Mineração de Dados Educacionais, Análise de Sentimentos e Ambientes Virtuais de Aprendizagem (fora do escopo) CE7. Estudos irrelevantes para a pesquisa, levando em consideração as questões de pesquisa CE8. Estudos que foram publicados antes de Janeiro de 2010

Fonte: elaborado pelos autores.

As etapas desse processo foram: (i) ler os títulos, resumos e palavras-chave dos artigos e, em seguida, excluir aqueles considerados irrelevantes para as questões de pesquisa; (ii) ler, na íntegra, os artigos selecionados na etapa anterior; (iii) documentar os artigos selecionados em um formulário predefinido.

2.4. Procedimentos de extração de dados

Os dados dos artigos selecionados foram extraídos de acordo com um formulário padrão predefinido que permitiu registrar os detalhes sobre os artigos. Assim, foram obtidos: título, ano de publicação, fonte, periódico/evento, autores, resumo e recursos discutidos. Além disso, todos os artigos foram categorizados com base na classificação de Zelkowitz e Wallace (1998), no qual foi considerado: (i) estudo de caso; (ii) experimental; (iii) lições aprendidas; e (iv) simulação.

Por fim, foi utilizada uma ferramenta web para apoiar revisões sistemáticas de literatura nominada de Parsifal² para apoiar a extração e o registro das informações dos estudos. Após o registro sobre os dados dos estudos foi aplicada uma análise quantitativa no sentido de verificar a distribuição dos artigos de acordo com a base de dados de literatura científica (Tabela 1) e o número de estudos primários publicados por ano (Figura 2). A partir disso, foram identificadas características e propriedades, relatadas nas Seções 3 e 4, de acordo com os objetivos e questões de pesquisa, propostos na Subseção 2.1.

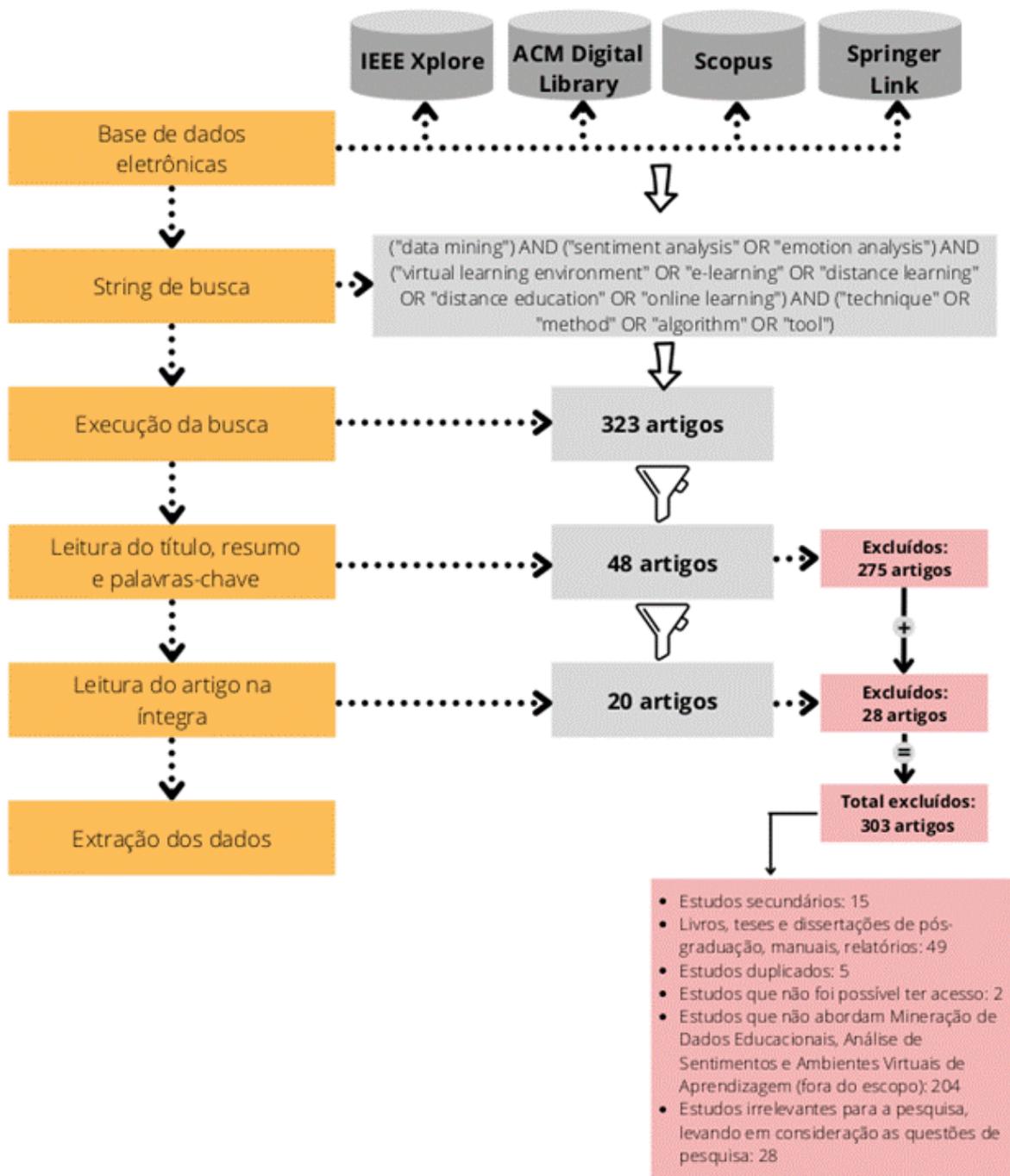
² <https://parsif.al>

3. Resultados

O mapeamento sistemático foi realizado durante um período de 90 dias no ano de 2021, de acordo com o protocolo apresentado na Seção 2. Após efetuados os procedimentos definidos nas Subseções 2.1, 2.2 e 2.3, um total de 20 artigos primários foram identificados, conforme apresentado na Figura 1.

Como mostrado na Figura 1, 323 artigos foram encontrados a partir da execução da *string* de busca. Em seguida, com base na leitura do título, resumo e palavras-chave, 48 artigos foram pré-selecionados. Na sequência, com a leitura na íntegra dos artigos, 20 estudos permaneceram para análise em profundidade.

Figura 1: Etapas realizadas no mapeamento sistemático.



Fonte: elaborada pelos autores.

A Tabela 1 apresenta a distribuição dos artigos de acordo com a base de dados de literatura científica.

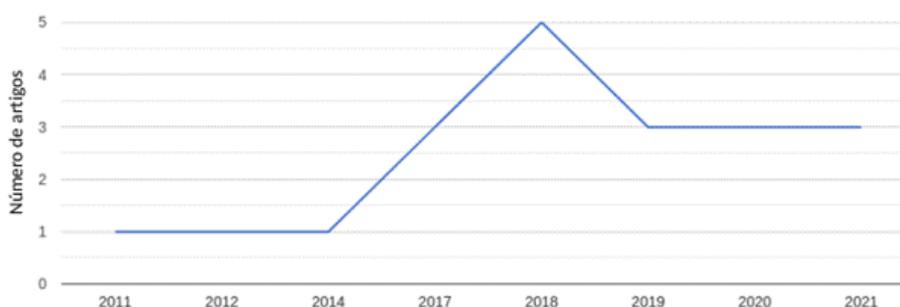
Tabela 1: Distribuição dos artigos encontrados de acordo com a fonte.

Base de dados de literatura científica	Artigos encontrados	Exclusão (pré-seleção) título + resumo + palavras-chave	Exclusão (seleção) leitura completa	Estudos primários identificados
ACM Digital Library	105	12	3	3
IEEE Xplore	7	3	3	3
Scopus	69	19	11	11
Springer Link	142	14	3	3
Total	323	48	20	20

Fonte: elaborada pelos autores.

Pode-se observar na Figura 2 que, no mapeamento sistemático realizado, não foram evidenciados muitos estudos primários sobre MDE e análise de sentimentos em AVA publicados até o ano de 2014. Além disso, o ano de 2018 teve o maior número de pesquisas publicadas, levando-se em consideração que este mapeamento sistemático da literatura foi concluído em novembro de 2021. Os resultados mostram um aumento nas pesquisas relacionadas à Mineração de Dados Educacionais e análise de sentimentos em Ambientes Virtuais de Aprendizagem, destacando sua importância na comunidade científica.

Figura 2: Número de estudos primários publicados por ano.



Fonte: elaborada pelos autores.

Os 20 estudos incluídos neste mapeamento são apresentados na Leitura Complementar, com números precedidos de EP (Estudo Primário) para distingui-los das referências. Além disso, foi elaborada uma planilha³ contendo os dados extraídos dos estudos primários selecionados. A seguir são apresentadas as características dos trabalhos encontrados por meio do mapeamento sistemático, de acordo com as questões de pesquisa definidas na Subseção 2.1.

QP1 - Quais são as técnicas, métodos, algoritmos, bibliotecas e ferramentas de mineração de dados educacionais para análise de base de dados referentes a sentimentos e emoções do estudante no processo de aprendizagem em AVA?

As técnicas, métodos, algoritmos, bibliotecas e ferramentas foram identificados nos 20 estudos primários, conforme apresentado no Quadro 5, correspondendo a QP1. Percebe-se uma predominância pelo uso de algoritmos de Aprendizado de Máquina, tais como: *Support Vector Machine*, *Naive Bayes* e *K-Nearest Neighbours* e pela ferramenta R para manipulação, análise e visualização dos dados sobre sentimentos. Nesse sentido, destaca-se o papel relevante do Aprendizado de Máquina por meio de tais algoritmos, pois se trata de uma técnica que trabalha com classificação de texto capaz de lidar com um conjunto de atribui-

3 <https://bit.ly/3oiuTAZ>

tos a serem analisados. Além disso, nota-se a importância que as ferramentas apresentam ao possibilitar uma análise de texto de forma automatizada a partir de um ambiente que apresenta funcionalidades para tal. No entanto, observa-se que não há uma biblioteca predominante para análise de dados referentes a sentimentos e emoções do estudante em AVA. Esse fato reforça a necessidade de estudos que realizem avaliações sobre as bibliotecas disponíveis para análise de sentimentos, apresentando suas características em relação ao seu desempenho, vantagens e desvantagens.

Quadro 5: Técnicas, métodos, algoritmos, bibliotecas e ferramentas identificadas nos estudos primários.

Procedimentos	Descrição	Quantidade	Estudos primários
Algoritmos	<i>Support Vector Machine (SVM)</i>	8	[EP01], [EP03], [EP04], [EP05], [EP06], [EP10], [EP15], [EP16]
	<i>Naive Bayes</i>	7	[EP02], [EP05], [EP06], [EP09], [EP10], [EP15], [EP16]
	<i>K-Nearest Neighbours (KNN)</i>	6	[EP03], [EP05], [EP06], [EP09], [EP14], [EP16]
	<i>Random Forest</i>	3	[EP05], [EP14], [EP16]
	<i>C4.5/J48 – Decision Tree</i>	2	[EP05], [EP09]
	<i>Logistic Regression</i>	2	[EP05], [EP16]
	<i>Hidden Model Markov (HMM)</i>	1	[EP04]
	Rede neural BiGRU e <i>Capsule Network</i>	1	[EP07]
	<i>Maximum Entropy</i>	1	[EP10]
	BERT-CNN	1	[EP13]
	<i>Additive Regression</i>	1	[EP14]
	<i>SMOreg (Sequential Minimal Optimization for Support Vector Regression)</i>	1	[EP14]
	<i>Boosting</i>	1	[EP15]
Bibliotecas	NLTK	2	[EP10], [EP12]
	CaTools do R	1	[EP01]
	NumPy	1	[EP10]
	SciPy	1	[EP10]
	Scikit-Learn	1	[EP10]
	Vader	1	[EP12]
	TextBlob	1	[EP12]
	Tensorflow	1	[EP16]
	Keras	1	[EP16]
Ferramentas	R	4	[EP01], [EP02], [EP17], [EP18]
	Weka	2	[EP03], [EP16]
	AntConc	1	[EP01]
	Rapid Miner	1	[EP09]
	Intellimote	1	[EP10]
	<i>Morphological analyser EAGLES</i>	1	[EP12]
	Word2Vec Gensim	1	[EP12]

Técnicas	word2vec	1	[EP06]
Outros	Linguagem de pré-treinamento ALBERT	2	[EP07], [EP17]
	Método baseado em <i>word segmentation, topic extraction</i> e construção de dicionário de polaridade	1	[EP08]
	Modelo não supervisionado ETJM (<i>Emotion Topic Joint Probabilistic Model</i>), extensão do modelo probabilístico Sentence-LDA (SLDA)	1	[EP11]
	Métodos de ensemble <i>learning: AdaBoost, Bagging, Random Subspace, Voting, Stacking</i>	1	[EP16]
	<i>Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Recurrent Neural Network with Attention Mechanism (RNN-AM), Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM)</i>	1	[EP16]
	Syuzhet (pacote R)	1	[EP17]
	NRC <i>Word-Emotion Association Lexicon</i>	1	[EP18]
	<i>Latent Dirichlet Allocation (LDA)</i>	1	[EP19]
	Classificação de polaridade de domínio cruzado	1	[EP20]
	<i>MeaningCloud</i>	1	[EP20]

Fonte: elaborado pelos autores.

QP2 - Com qual objetivo foi empregada a análise de sentimentos em AVA?

Os motivos pelos quais os pesquisadores analisaram sentimentos em AVA foram identificados nos 20 estudos primários, conforme apresentado no Quadro 6, correspondendo a QP2. Cabe ressaltar que, os pesquisadores empregam a análise de sentimentos com diferentes finalidades: avaliar cursos e professores, verificar a eficácia do AVA para melhorar o processo de ensino e, também, fornecer aprendizagem personalizada. Contudo, observa-se a predominância por investigar as opiniões dos alunos sobre cursos, professores e material didático. Tal fato reforça a importância que a análise de sentimentos em AVA apresenta quando se trata da qualidade dos materiais didáticos no sucesso dos estudantes durante o processo de aprendizado. Além disso, a análise de sentimentos dos estudantes evidencia como a formação contínua dos professores e a atualização dos cursos influenciam na motivação e engajamento dos estudantes durante o aprendizado. Entretanto, diante das informações apresentadas no Quadro 6, observa-se que ainda são necessários estudos sobre a análise de sentimentos em AVA para a previsão da desistência de estudantes no curso e formação de equipes de trabalho. Assim, a partir da identificação do estado emocional dos indivíduos, será possível aplicar estratégias pedagógicas que aumentem o nível de confiança dos estudantes para alcançar sucesso no processo de aprendizagem.

Quadro 6: Objetivos para análise de sentimentos em AVA identificados nos estudos primários.

Categoria	Quantidade	Estudos primários	Objetivos
Investigar as opiniões dos alunos sobre cursos, professores e material didático.	7	[EP01]	Analisar as opiniões dos alunos sobre cursos em um MOOC (<i>Massive Open Online Course</i>), seus instrutores e as principais ferramentas utilizadas no curso a partir de dados coletados de e-mails trocados entre estudantes de MOOC.
		[EP08]	Avaliar professores por meio de comentários de alunos no fórum do AVA para formar um repositório de avaliação de professores.
		[EP09]	Classificar e prever a qualidade do feedback afetivo para aluno com base no conteúdo do <i>feedback</i> e no estado afetivo do aluno em um fórum de aprendizagem de inglês.
		[EP12]	Extrair e analisar a opinião de estudantes sobre o material didático a partir de fóruns em MOOCs, para melhorar os materiais dos cursos online, obter percepções sobre serviços de aprendizagem, detectar os usuários que atrapalham o trabalho em grupo.
		[EP13]	Avaliar cursos a partir de comentários em fórum de MOOCs.
		[EP15]	Prever aspectos essenciais (conteúdo, estrutura, instrutor, design do curso) que contribuem para a satisfação dos alunos na aprendizagem online a partir das avaliações dos alunos e, em seguida, avaliar a atitude dos alunos em relação a esses aspectos.
		[EP20]	Avaliar experiência do usuário em ambiente virtual de aprendizagem (Conecto e Moodle) por meio de comentários de usuários.
Melhorar o processo de ensino.	4	[EP03]	Classificar as opiniões dos alunos de acordo com sua polaridade e analisá-las a fim de contribuir para a melhoria do processo de ensino em uma plataforma de treinamento educacional.
		[EP11]	Explorar opiniões negativas de comentários em MOOC para que professores possam regulamentar e melhorar os métodos de ensino, estratégias e conteúdos de aprendizagem.
		[EP16]	Obter <i>feedback</i> sobre o conteúdo de aprendizagem, analisando avaliações de MOOCs, o que pode auxiliar professores a melhorar seu processo de ensino, por meio de um esquema de classificação de sentimento com desempenho preditivo aplicando paradigmas de <i>ensemble learning</i> e <i>deep learning</i> .
		[EP18]	Compreender os problemas de ensino e aprendizagem a partir dos comentários dos alunos em formulários.
Fornecer aprendizagem personalizada.	3	[EP06]	Analisar a mudança de emoção durante aprendizagem dos alunos e fornecer aprendizagem personalizada aos alunos, a partir dos dados de interação em fórum do MOOC.
		[EP10]	Analisar comentários de estudantes sobre materiais de <i>e-learning</i> por meio de análise de sentimento para selecionar objetos de aprendizagem para serem reutilizados.
		[EP19]	Sugerir novos conteúdos personalizados para as aulas ou caminhos de aprendizado adaptados aos alunos, a partir da análise de sentimentos de postagens coletadas no Moodle.

Melhorar o desempenho dos alunos.	2	[EP04]	Classificar a opinião do aluno a partir de fóruns e blogs de <i>e-learning</i> para melhorar seu desempenho.
		[EP05]	Analisar a satisfação e dificuldades dos alunos a partir de fóruns de discussão retirados de mais de cinco anos de um curso universitário de Introdução à Programação em Java, ministrado por meio do AVA Moodle.
Avaliar a eficácia e qualidade do AVA.	2	[EP02]	Apresentar a polaridade dos textos e a caracterização das emoções a partir da participação de alunos no fórum da plataforma Moodle de um curso de pós-graduação de uma universidade com o objetivo de avaliar a eficácia do ambiente de aprendizagem para melhorar a experiência de aprendizagem dos alunos, a experiência de ensino dos tutores e a visão estratégica institucional da universidade.
		[EP07]	Obter uma compreensão aprofundada a partir de comentários de avaliação de curso MOOC para melhorar a qualidade da plataforma.
Prever desistência do aluno no curso.	1	[EP14]	Prever os alunos que estão em risco de abandono ou precisam de suporte adicional para aumentar seu sucesso e confiança no processo de aprendizagem, por meio da previsão de notas, levando em consideração as mensagens no Moodle e a mudança do estado emocional dos usuários. Promover a autoavaliação dos tutores e intervenções focadas para melhorar as aulas e aumentar a qualidade do suporte dos alunos.
Formar equipes de trabalho.	1	[EP17]	Ajudar o professor a identificar os níveis de motivação e possíveis problemas de seus alunos diante do curso. Servir como guia do professor para a formação de equipes de trabalho, atribuição de papéis em projetos

Fonte: elaborado pelos autores.

QP3 - Quais tipos de emoções e sentimentos têm sido considerados nos estudos?

Os tipos de emoções e sentimentos foram identificados nos 20 estudos primários, conforme apresentado no Quadro 7, correspondendo a QP3. A maioria das pesquisas aponta apenas a polaridade de sentimentos: positiva, negativa e neutra. Porém, um artigo aborda outras escalas para a polaridade de sentimentos: muito negativo, negativo, neutro, positivo e muito positivo. Apenas quatro estudos exploram emoções na mineração de dados educacionais em AVA. As emoções mais comuns são: raiva, nojo, medo, alegria, tristeza e surpresa.

Quadro 7: Tipos de emoções e sentimentos identificados nos estudos primários.

Emoções e sentimentos	Quantidade	Estudos primários
Polaridade (positiva e negativa) de sentimentos	8	[EP01], [EP04], [EP06], [EP08], [EP13], [EP14], [EP16], [EP19]
Polaridade (positiva, negativa, neutra) de sentimentos	7	[EP02], [EP07], [EP10], [EP12], [EP15], [EP18], [EP20]
Polaridade (muito negativo, negativo, neutro, positivo, muito positivo) de sentimentos	1	[EP03]
Oito emoções (raiva, antecipação/expectativa, nojo, medo, alegria, tristeza, surpresa, confiança)	3	[EP14], [EP17], [EP18]

Seis emoções (raiva, nojo, medo, alegria, tristeza e surpresa)	1	[EP02]
Estados afetivos (frustrado, desmotivado, ansioso e entediado)	1	[EP09]
Apenas emoções negativas	1	[EP11]
Não deixam explícitas quais seriam as emoções; apenas citam frustração e confusão.	1	[EP05]

Fonte: elaborado pelos autores.

4. Discussão

A partir dos resultados encontrados nos estudos primários, verifica-se que análise de sentimentos em AVA é empregada para: (1) investigar as opiniões dos alunos sobre cursos, professores e material didático (EP01, EP08, EP09, EP12, EP13, EP15, EP20); (2) melhorar o processo de ensino (EP03, EP11, EP16, EP18); (3) fornecer aprendizagem personalizada (EP06, EP10, EP19); (4) melhorar o desempenho dos alunos (EP04, EP05); (5) avaliar a eficácia e qualidade do AVA (EP02, EP07); (6) prever desistência do aluno no curso (EP14); (7) formar equipes de trabalho (EP17).

A fim de obter uma visão aprofundada sobre a análise de sentimentos durante o processo de ensino e aprendizagem, o foco deste mapeamento sistemático consistiu em AVA. Assim, é possível identificar lacunas que podem ser exploradas no futuro em pesquisas sobre MDE e análise de sentimentos nesses ambientes. Algumas das questões em aberto mais importantes em diferentes aspectos da análise de sentimento são apresentadas a seguir.

A partir dos estudos primários encontrados neste mapeamento sistemático, percebe-se uma tendência por algoritmos de aprendizado de máquina, tais como, SVM, *Naive Bayes* e KNN, que identificam emoções e sentimentos a partir das interações dos alunos, principalmente, por meio de fóruns. Tal predominância se deve ao menor grau de dificuldade em capturar o formato de texto. No entanto, para que mecanismos em AVA possam analisar emoções e sentimentos em cenários educacionais da realidade, é necessário que pesquisas futuras se concentrem em dados produzidos em ambientes de ensino e aprendizagem do mundo real. Tais dados dizem respeito aos comportamentos e interações dos alunos em diferentes momentos na utilização do AVA, sejam síncronos ou assíncronos, tais como: em atividades em grupo, respondendo uma avaliação ou procurando solucionar dúvidas de conteúdo.

Os estudos primários encontrados não abordam a granularidade da representação do sentimento e/ou emoção dos alunos em AVA. Se o ambiente permitir que o professor defina a periodicidade do *feedback* sobre os sentimentos dos alunos em semanas, diariamente, ou até mesmo em horas ou minutos, é possível acompanhar o processo constantemente para descobrir eventuais deficiências de ensino e aprendizagem.

Aponta-se a predominância de estudos primários que exploram apenas a polaridade de sentimentos (positiva, negativa e neutra). No entanto, poucos pesquisadores abordam emoções, tais como: raiva, nojo, medo, alegria, tristeza e surpresa. Tais emoções podem manifestar detalhes sobre aspectos que interferem no desenvolvimento cognitivo e social dos alunos. Nesse sentido, as pesquisas deveriam dar maior enfoque na análise de emoções em AVA e, conseqüentemente, melhorar o desempenho na verificação de eventuais obstáculos durante o processo de aprendizagem ou na descoberta da relação entre a atuação de um aluno e o seu interesse pelo curso.

Destaca-se que, ao explorar outras fontes de emoções, tais como as expressões faciais, é viável realizar uma análise aprimorada sobre os sentimentos dos alunos. Dessa forma, é possível compreender como as

relações de afeto são construídas no ambiente de aprendizagem e definir quais estratégias pedagógicas o professor poderia colocar em prática. Além disso, é importante verificar os sentimentos e as emoções também dos professores em AVA, o que pode ser significativo para a gestão da instituição educacional.

5. Conclusão

O presente artigo apresenta um mapeamento sistemático da literatura sobre MDE e Análise de Sentimentos em AVA. Para tal, foi relatado o processo de coleta, avaliação e sistematização de estudos primários por meio de uma abordagem mais ampla e quantitativa, fornecendo dados para identificação de tendências de estudos, bem como lacunas a serem exploradas (PETERSEN *et al.*, 2015). Assim, é possível fornecer apoio para pesquisas futuras focadas em MDE, além de análise de sentimentos e AVA.

Em relação aos resultados, foram evidenciados dados quantitativos, incluindo o número de estudos por etapas de seleção para cada base de dados de literatura científica e ano. Em seguida, foram sintetizados os estudos primários e discutidas algumas características dos trabalhos, de acordo com as questões de pesquisa. Assim, foi possível obter uma visão geral sobre a literatura científica atual em MDE, análise de sentimentos e AVA. Enfatiza-se que a busca foi reduzida a um número limitado de buscadores e foram excluídos os estudos sobre análise de sentimentos que não eram aplicados em AVA ou que não contribuíram com nenhum método significativo no contexto deste mapeamento sistemático da literatura.

Compreende-se que é necessário explorar outras fontes de informações para capturar sentimentos e emoções dos alunos, para além de fóruns, conforme abordado nos estudos primários encontrados. Como próximos passos, pretende-se realizar um mapeamento sistemático sobre relatos de experiências de análise de sentimentos e emoções em AVA que enfatizem o período de pandemia causado pelo COVID-19.

É importante destacar também que, redes sociais vinculadas ao ambiente de aprendizagem, por meio do cruzamento de informações, podem viabilizar análises de sentimentos dos alunos mais sofisticadas e próximas da realidade. Além disso, para a continuidade de pesquisas na área de estudo abordada neste mapeamento sistemático, sugere-se a realização de estudos empíricos comparando o desempenho de várias técnicas, algoritmos e bibliotecas de análise de sentimentos aplicados em AVA. Espera-se que os resultados divulgados neste artigo contribuam para pesquisas futuras sobre Educação a Distância, principalmente em investigações com ênfase nos efeitos das emoções e sentimentos sobre o desempenho em atividades e na conclusão dos cursos. Além disso, o estudo permite discutir a recomendação de estratégias adequadas com o perfil socioafetivo dos estudantes em AVA e agir antecipadamente mediante a detecção de sinais comportamentais indicativos de abandono do curso.

Referências

- ANITHA, N.; ANITHA, B.; PRADEEPA, S. Sentiment classification approaches. **International Journal of Innovation Engineering and Technology**, v. 3, n. 1, p. 22-31, 2013.
- ARUNA, S.; SASANKA, J.; VINAY, D. A. A brief study on analyzing student's emotions with the help of educational data mining. **Computer Networks, Big Data and IoT**, p. 785-796, 2021.
- HEUBNER, R. A. A survey of educational data-mining research. **Research in Higher Education Journal**, v. 19, 2013.
- PETERSEN, K.; VAKKALANKA, S.; KUZNIARZ, L. Guidelines for conducting systematic mapping studies in software engineering: an update. **Information and Software Technology**, v. 64, p. 1-18, 2015.
- PURAIWAN, E. *et al.* Emotion-based decision support tool for learning processes: an application with under-

graduate students during Covid-19 pandemic. In: **Iberian Conference on Information Systems and Technologies**. 2021. p. 1-6.

SACHIN, R. B.; VIJAY, M. S. A survey and future vision of data mining in educational field. In: **International Conference on Advanced Computing & Communication Technologies**. 2012. p. 96-100.

ZELKOWITZ, M. V.; WALLACE, D. R. Experimental models for validating technology. **Computer**, v. 31, n. 5, p. 23-31, 1998.

Leitura Complementar

[EP01] BUENAÑO-FERNÁNDEZ, D.; LUJÁN-MORA, S.; VILLEGAS-CH, W. Improvement of massive open online courses by text mining of students' emails: a case study. In: **International Conference on Technological Ecosystems for Enhancing Multiculturality**. 2017. p. 1-7.

[EP02] GKONTZIS, A. F. *et al.* Sentiment analysis to track emotion and polarity in student fora. In: **Pan-Hellenic Conference on Informatics**. 2017. p. 1-6.

[EP03] SPATIOTIS, N. *et al.* Evaluation of an educational training platform using text mining. In: **Hellenic Conference on Artificial Intelligence**. 2018. p. 1-5.

[EP04] KECHAOU, Z.; AMMAR, M. B.; ALIM, A. M. Improving e-learning with sentiment analysis of users' opinions. In: **EDUCON IEEE Global Engineering Education Conference**. 2011. p. 1032-1038.

[EP05] HARRIS, S. C.; KUMAR, V. Identifying student difficulty in a digital learning environment. In: **ICALT International Conference on Advanced Learning Technologies**. 2018. p. 199-201.

[EP06] FEI, H.; LI, H. The study of learners' emotional analysis based on MOOC. In: **International Conference on Cognitive Computing**. 2018. p. 170-178.

[EP07] LIU, T. *et al.* Sentiment analysis for MOOC course reviews. In: **International Conference of Pioneering Computer Scientists, Engineers and Educators**. Springer, 2021. p. 78-87.

[EP08] CAIQIANG, L.; JUNMING, M. Research on online education teacher evaluation model based on opinion mining. In: **CITCS National Conference on Information Technology and Computer Science**. 2012.

[EP09] SELMI, M.; HAGE, H.; AÏMEUR, E. Opinion Mining for predicting peer affective feedback helpfulness. In: **KMIS**. 2014. p. 419-425.

[EP10] MANDAL, L.; DAS, R.; BHATTACHARYA, S.; BASU, P. N. Intellimote: a hybrid classifier for classifying learners' emotion in a distributed e-learning environment. **Turkish Journal of Electrical Engineering & Computer Sciences**, v. 25, n. 3, p. 2084-2095, 2017.

[EP11] LIU, Z. *et al.* Joint exploration of negative academic emotion and topics in student-generated online course comments. In: **EITT International Conference of Educational Innovation through Technology**. IEEE, 2017. p. 89-93.

[EP12] COBOS, R.; JURADO, F.; VILLÉN, Á. Moods in MOOCs: analyzing emotions in the content of online courses with edX-CAS. In: **EDUCON Global Engineering Education Conference**. IEEE, 2019. p. 1467-1474.

[EP13] LI, X. *et al.* A Shallow BERT-CNN Model for Sentiment Analysis on MOOCs Comments. In: **TALE International Conference on Engineering, Technology and Education**. IEEE, 2019. p. 1-6.

[EP14] GKONTZIS, A. F. *et al.* Polarity, emotions and online activity of students and tutors as features in predicting grades. **Intelligent Decision Technologies**, v. 14, n. 3, p. 409-436, 2020.

[EP15] KASTRATI, Z. *et al.* Aspect-based opinion mining of students' reviews on online courses. In: **International Conference on Computing and Artificial Intelligence**. 2020. p. 510-514.

[EP16] ONAN, A. Sentiment analysis on massive open online course evaluations: a text mining and deep learning approach. **Computer Applications in Engineering Education**, v. 29, n. 3, p. 572-589, 2021.

[EP17] PURAIVAN, E. *et al.* Emotion-based decision support tool for learning processes: an application with undergraduate students during Covid-19 pandemic. In: **CISTI Iberian Conference on Information Systems and Technologies**. IEEE, 2021. p. 1-6.

[EP18] PRAVEENKUMAR, T.; MANORSELVI, A.; SOUNDARAPANDIYAN, K. Exploring the students feelings and emotion towards online teaching: sentimental analysis approach. In: **International Working Conference on Transfer and Diffusion of IT**. Springer, 2020. p. 137-146.

[EP19] CLARIZIA, F. *et al.* E-learning and sentiment analysis: a case study. In: **International Conference on Information and Education Technology**. 2018. p. 111-118.

[EP20] SANCHIS-FONT, R.; CASTRO-BLEDA, M. J.; GONZÁLEZ, J. Á. Applying sentiment analysis with cross-domain models to evaluate user eXperience in virtual learning environments. In: **International Work-Conference on Artificial Neural Networks**. Springer, 2019. p. 609-620.