Review

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

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

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

Rafael Leonardo Vivian[1,2*]
Silvio César Cazella[3]
Leticia Rocha Machado[1]
Patricia Alejandra Behar[1]

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

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

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

*rafael.vivian@ifc.edu.br

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.

Received:: 04/06/2022
Accepted: 06/14/2022
Published: 06/22/2022

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

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

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.*

1. Introduction

The increase in the use of information and communication technologies in educational systems has provided the storage of large amounts of student data, which allows for the use of Educational Data Mining (EDM) to improve teaching and learning processes (HEUBNER, 2013); SACHIN and VIJAY, 2012). EDM is useful in several areas, including identifying feelings and emotions, discovering learning needs for different groups of learners (ARUNA, SASANKA and VINAY, 2021). In this way, sentiment analysis can reveal a relationship between learning performance and student interest in a particular course.

Sentiment analysis is understood as an opinion mining technique and can involve: Natural Language Processing (NLP), information extraction, Artificial Intelligence, Machine Learning and Data Mining (PURAIVAN et al., 2021). Thus, this technique is divided into four phases: data acquisition, data preparation, review analysis and sentiment classification. Furthermore, there are two main approaches: (1) machine learning, which is divided into supervised and unsupervised; and (2) lexicon-based approach, which is divided into two categories, dictionary-based and corpus-based (ANITHA, ANITHA and PRADEEPA, 2013).

In Virtual Learning Environments (VLE), interactions between students, teacher and tutor occur mainly through messages in forums and chats. The different subjects and the different contents of the courses can generate an amount of information in the VLE that makes it difficult for teachers and tutors to examine the texts manually to check the students' feelings and to discover possible obstacles during the teaching and learning process.

In this scenario, the objective of this article is to present the results of a systematic mapping of the literature on EDM and Sentiment Analysis in VLE. Thus, this systematic mapping made it possible to identify the main aspects of the literature on the subject from 2010 to 2021, published in scientific journals and events, in addition to verifying opportunities for future work.

This article is organized as follows. Section 2 deals with the systematic mapping procedure applied; Section 3 describes the results obtained; Section 4 presents the discussion and, finally, Section 5 closes with conclusions and limitations.

2. Methodology

Systematic literature mapping is a process of collecting, evaluating and systematizing primary studies. Thus, a broader and more quantitative approach is presented that provides data to identify trends in studies, as well as gaps to be explored (PETERSEN et al., 2015). From this perspective, the focus of this systematic literature mapping is to identify primary studies on EDM and Sentiment Analysis in VLE. Therefore, this type of study was chosen to obtain an overview of the areas, with regard to the techniques, methods, algorithms, libraries and tools used; the purposes of sentiment analysis; and the emotions and feelings analyzed.

2.1. Research questions

The research questions (RQ) were prepared considering the outlining objective, being presented in Chart 1, in order to verify the evidence in the scientific literature on EDM, sentiment analysis and VLE.

Chart 1: Research questions.

RQ1	What are the educational data mining techniques, methods, algorithms, libraries and tools for database analysis regarding student feelings and emotions in the learning process in VLE?
RQ2	For what purpose was sentiment analysis used in VLE?
RQ3	What types of emotions and feelings have been considered in studies?

Source: designed by the authors

2.2. Search process

The research was carried out in scientific literature databases using English keywords for the categories “data mining”, “sentiment analysis” and “virtual learning environments”, as shown in Chart 2. In each category, match the keywords using the Boolean operator “OR”; both categories were grouped using the Boolean operator “AND”. The following scientific literature databases were searched:

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)

Chart 2: Keywords used in this mapping.

Reference	Category	Keywords
C1	Data mining	Data mining
C2	Sentiment analysis	Sentiment analysis
		Emotion analysis
C3	Virtual learning environment	Virtual learning environment
		E-learning
		Distance learning
		Distance education
		Online learning
C4	Procedures	Technique
		Method
		Algorithm
		Tool

Source: designed by the authors

Thus, full scientific articles focused on “data mining”, “sentiment analysis” and “virtual learning environment”, written in English, published between January 2010 and September 2021 were searched. The search string was defined as a combination of C1, C2, C3 and C4, using the Boolean operators “AND” and “OR”, as shown in Chart 3.

Chart 3: *Search string applied to each database*.

Base de dados	String
ACM Digital Library	[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”]]
IEEE Xplore	(“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”)
Scopus	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”))
Springer Link	(“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”)

Source: designed by the authors

The search string was adapted according to the syntax of each scientific literature database, since the search options differ and the way it should be constructed is specific to each database.

2.3. Inclusion and exclusion criteria

The inclusion and exclusion criteria used in this systematic mapping applied in the research are those listed in Chart 4.

Chart 4: Inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria
IC1. primary studies IC2. Studies that address Educational Data Mining, Sentiment Analysis and Virtual Learning Environments	EC1. secondary studies EC2. Books, book chapters, graduate theses and dissertations, manuals, reports EC3. Duplicate studies (only one copy of each study was included) EC4. Studies that could not be accessed EC5. Studies not written in English EC6. Studies that do not address Educational Data Mining, Sentiment Analysis and Virtual Learning Environments (out of scope) EC7. Studies irrelevant to the research, taking into account the research questions EC8. Studies that were published before January 2010

Source: designed by the authors

The steps in this process were: (i) reading the titles, abstracts and keywords of the articles and then excluding those considered irrelevant to the research questions; (ii) read, in full, the articles selected in the previous step; and (iii) document selected articles in a predefined form.

2.4. Data extraction procedures

Data from the selected articles were extracted according to a predefined standard form that allowed recording the details about the articles. Thus, the following were obtained: title, year of publication, source, journal/event, authors, abstract and discussed resources. In addition, all articles were categorized based on the classification of Zelkowitz and Wallace (1998), which considered: (i) case study; (ii) experimental; (iii) lessons learned; and (iv) simulation.

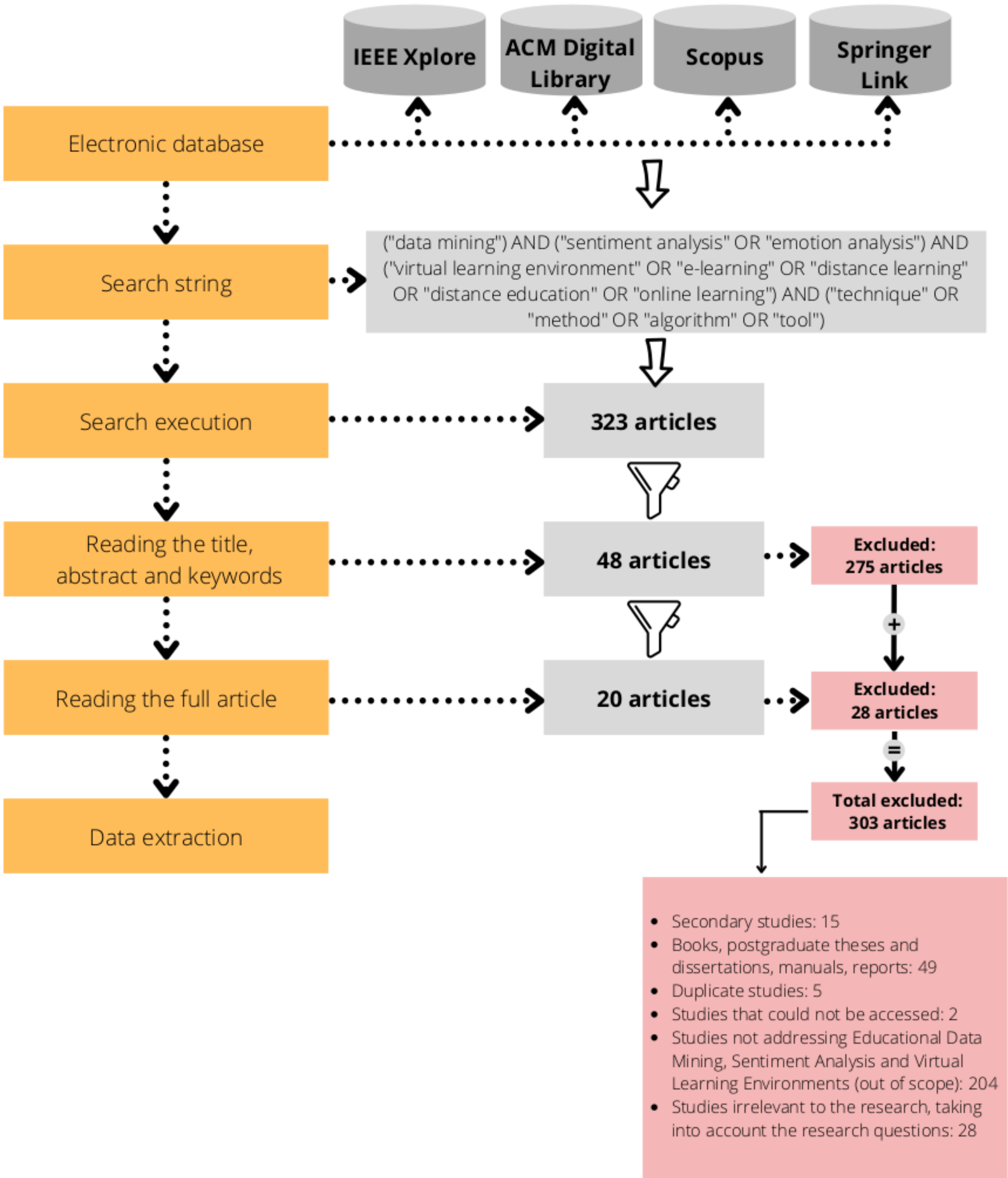
Finally, a web tool was used to support systematic reviews of literature named by Parsifal to support the extraction and recording of information from the studies. After recording the data from the studies, a quantitative analysis was applied to verify the distribution of articles according to the scientific literature database (Table 1) and the number of primary studies published per year (Figure 2). From this, characteristics and properties were identified, reported in Sections 3 and 4, according to the research objectives and questions proposed in Subsection 2.1.

3. Results

The systematic mapping was carried out over a period of 90 days in the year 2021, according to the protocol presented in Section 2. After carrying out the procedures defined in Subsections 2.1, 2.2 and 2.3, a total of 20 primary articles were identified, as presented in Figure 1.

As shown in Figure 1, 323 articles were found from the execution of the search string. Then, based on the reading of the title, abstract and keywords, 48 articles were pre-selected. Subsequently, with the full reading of the articles, 20 studies remained for in-depth analysis.

Figure 1: Steps performed in the systematic mapping.



Source: designed by the authors

Table 1 shows the distribution of articles according to the scientific literature database.

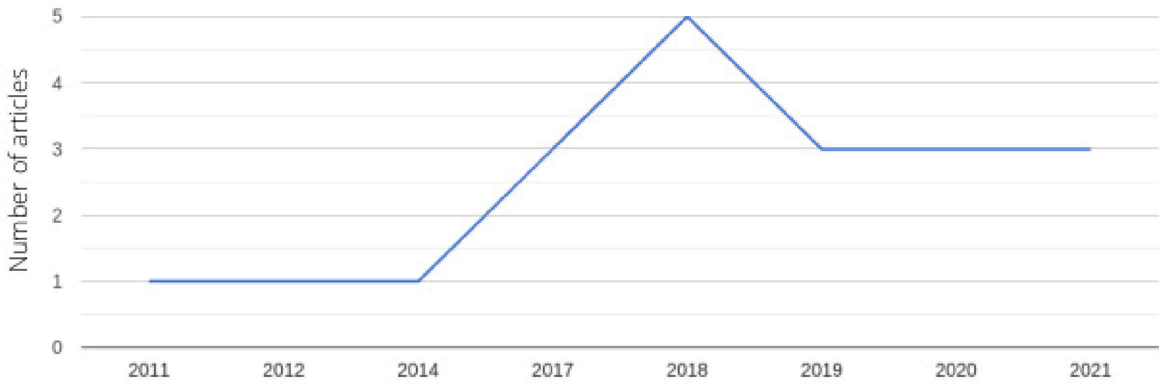
Table 1: Distribution of articles found according to source.

Scientific literature database	Articles found	Exclusion (pre-selection) title + abstract + keywords	Deletion (selection) full reading	Primary studies identified
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

Source: designed by the authors

It can be seen in Figure 2 that, in the systematic mapping carried out, many primary studies on EDM and sentiment analysis in VLE published until 2014 were not evidenced. In addition, the year 2018 had the highest number of published studies, taking into account that this systematic literature mapping was completed in November 2021. The results show an increase in research related to Educational Data Mining and sentiment analysis in Virtual Learning Environments, highlighting its importance in the scientific community.

Figure 2: Number of primary studies published per year.



Source: designed by the authors

The 20 studies included in this mapping are presented in the Complementary Reading, with numbers preceded by PS (Primary Study) to distinguish them from the references. In addition, a spreadsheet was prepared containing the data extracted from the selected primary studies. The characteristics of the works found through systematic mapping are presented below, according to the research questions defined in Subsection 2.1.

RQ1 - What are the educational data mining techniques, methods, algorithms, libraries and tools for database analysis regarding student feelings and emotions in the learning process in VLE?

Techniques, methods, algorithms, libraries and tools were identified in the 20 primary studies, as shown in Chart 5, corresponding to RQ1. There is a predominance of the use of Machine Learning algorithms, such as: Support Vector Machine, Naive Bayes and K-Nearest Neighbors and the R tool for manipulation, analysis and visualization of sentiment data. In this sense, the relevant role of Machine Learning through such algorithms is highlighted, as it is a technique that works with text classification capable of dealing with a set of attributes to be analyzed. Furthermore, it is noted the importance that the tools present in enabling an automated text analysis from an environment that has functionalities for this purpose. However, it is observed that there is no predominant library for data analysis related to student feelings and emotions in VLE. This fact reinforces the need for studies that carry out evaluations on the libraries available for sentiment analysis, presenting their characteristics in relation to their performance, advantages and disadvantages.

Chart 5: Techniques, methods, algorithms, libraries and tools identified in primary studies.

Procedures	Description	Quantity	Primary Studies
Algorithms	Support Vector Machine (SVM)	8	[PS01], [PS03], [PS04], [PS05], [PS06], [PS10], [PS15], [PS16]
	Naive Bayes	7	[PS02], [PS05], [PS06], [PS09], [PS10], [PS15], [PS16]
	K-Nearest Neighbours (KNN)	6	[PS03], [PS05], [PS06], [PS09], [PS14], [PS16]
	Random Forest	3	[PS05], [PS14], [PS16]
	C4.5/J48 – Decision Tree	2	[PS05], [PS09]
	Logistic Regression	2	[PS05], [PS16]
	Hidden Model Markov (HMM)	1	[PS04]
	Rede neural BiGRU e Capsule Network	1	[PS07]
	Maximum Entropy	1	[PS10]
	BERT-CNN	1	[PS13]
	Additive Regression	1	[PS14]
	SMOreg (Sequential Minimal Optimization for Support Vector Regression)	1	[PS14]
	Boosting	1	[PS15]
Libraries	NLTK	2	[PS10], [PS12]
	CaTools do R	1	[PS01]
	NumPy	1	[PS10]
	SciPy	1	[PS10]
	Scikit-Learn	1	[PS10]
	Vader	1	[PS12]
	TextBlob	1	[PS12]
	Tensorflow	1	[PS16]
	Keras	1	[PS16]
Tools	R	4	[PS01], [PS02], [PS17], [PS18]
	Weka	2	[PS03], [PS16]
	AntConc	1	[PS01]
	Rapid Miner	1	[PS09]
	Intellimote	1	[PS10]
	Morphological analyser EAGLES	1	[PS12]
	Word2Vec Gensim	1	[PS12]
Techniques	word2vec	1	[PS06]

Others	ALBERT pre-training language	2	[PS07], [PS17]
	Method based on word segmentation, topic extraction and polarity dictionary construction	1	[PS08]
	unsupervised model ETJM (Emotion Topic Joint Probabilistic Model), extension of the Sentence-LDA probabilistic model (SLDA)	1	[PS11]
	Ensemble learning methods: AdaBoost, Bagging, Random Subspace, Voting, Stacking	1	[PS16]
	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)	1	[PS16]
	Syuzhet (pacote R)	1	[PS17]
	NRC Word-Emotion Association Lexicon	1	[PS18]
	Latent Dirichlet Allocation (LDA)	1	[PS19]
	Cross Domain Polarity Classification	1	[PS20]
	MeaningCloud	1	[PS20]

Source: designed by the authors

RQ2 - For what purpose was sentiment analysis used in VLE?

The reasons why the researchers analyzed feelings in VLE were identified in the 20 primary studies, as shown in Chart 6, corresponding to RQ2. It should be noted that researchers use sentiment analysis for different purposes: evaluating courses and teachers, verifying the effectiveness of VLE to improve the teaching process and also providing personalized learning. However, there is a predominance of investigating students' opinions about courses, teachers and didactic material. This fact reinforces the importance that sentiment analysis in VLE presents when it comes to the quality of teaching materials in the success of students during the learning process. In addition, the analysis of students' feelings shows how the continuous training of teachers and the updating of courses influence students' motivation and engagement during learning. However, in view of the information presented in Chart 6, it is observed that studies on the analysis of feelings in VLE are still needed to predict the dropout of students in the course and the formation of work teams. Thus, from the identification of the emotional state of individuals, it will be possible to apply pedagogical strategies that increase the students' confidence level to achieve success in the learning process.

Chart 6: Objectives for sentiment analysis in VLE identified in primary studies.

Category	Quantity	Primary Studies	Objectives
Investigate student opinions about courses, teachers and teaching materials	7	[PS01]	Analyze students' opinions about courses in a MOOC (Massive Open Online Course), their instructors and the main tools used in the course from data collected from emails exchanged between MOOC students.
		[PS08]	Evaluate teachers through student feedback on the VLE forum to form a teacher assessment repository.
		[PS09]	Rate and predict the quality of affective student feedback based on feedback content and student affective state in an English learning forum.
		[PS12]	Extract and analyze student opinion about teaching material from forums in MOOCs, to improve online course materials, gain insights into learning services, detect users that interfere with group work.
		[PS13]	Evaluate courses from MOOC forum comments.
		[PS15]	Predict essential aspects (content, structure, instructor, course design) that contribute to student satisfaction in online learning from student assessments and then assess students' attitude towards these aspects.
		[PS20]	Evaluate user experience in a virtual learning environment (Conecto and Moodle) through user comments.
Improve the teaching process	4	[PS03]	Classify students' opinions according to their polarity and analyze them in order to contribute to the improvement of the teaching process in an educational training platform.
		[PS11]	Explore negative opinions of MOOC comments so that teachers can regulate and improve teaching methods, strategies and learning content.
		[PS16]	Get feedback on learning content by analyzing MOOC assessments, which can help teachers improve their teaching process, through a sentiment rating scheme with predictive performance applying ensemble learning and deep learning paradigms.
		[PS18]	Understand teaching and learning problems from students' comments on forms.
Provide personalized learning	3	[PS06]	Analyze changing emotion during student learning and provide personalized learning to students from MOOC forum interaction data.
		[PS10]	Analyze student feedback on e-learning materials through sentiment analysis to select learning objects to be reused.
		[PS19]	Suggest new personalized content for classes or learning paths adapted to students, based on sentiment analysis of posts collected in Moodle.
Improve student performance	2	[PS04]	Grade student opinion from e-learning forums and blogs to improve their performance.
		[PS05]	To analyze the satisfaction and difficulties of students from discussion forums taken from more than five years of a university course on Introduction to Programming in Java, taught through the VLE Moodle.

Evaluate the effectiveness and quality of the VLE	2	[PS02]	To present the polarity of texts and the characterization of emotions from the participation of students in the Moodle platform forum of a postgraduate course at a university with the objective of evaluating the effectiveness of the learning environment to improve the students' learning experience , the tutors' teaching experience and the university's institutional strategic vision.
		[PS07]	Gain in-depth understanding from MOOC course evaluation comments to improve platform quality.
Predict student dropout from the course	1	[PS14]	Predict learners who are at risk of dropping out or need additional support to increase their success and confidence in the learning process, through grade prediction, taking into account Moodle messages and changing emotional state of users. Promote tutor self-assessment and focused interventions to improve classes and increase the quality of student support.
Form work teams	1	[PS17]	Help the teacher to identify the levels of motivation and possible problems of their students in the course. Serve as a teacher's guide for forming work teams, assigning roles in projects

Source: designed by the authors

RQ3 - What types of emotions and feelings have been considered in the studies?

The types of emotions and feelings were identified in the 20 primary studies, as shown in Chart 7, corresponding to RQ3. Most research points only to the polarity of feelings: positive, negative and neutral. However, an article discusses other scales for the polarity of feelings: very negative, negative, neutral, positive and very positive. Only four studies explore emotions in educational data mining in VLE. The most common emotions are: anger, disgust, fear, joy, sadness and surprise.

Chart 7: Types of emotions and feelings identified in primary studies.

Emotions and feelings	Quantity	Primary studies
Polarity (positive and negative) of feelings	8	[PS01], [PS04], [PS06], [PS08], [PS13], [PS14], [PS16], [PS19]
Polarity (positive, negative, neutral) of feelings	7	[PS02], [PS07], [PS10], [PS12], [PS15], [PS18], [PS20]
Polarity (very negative, negative, neutral, positive, very positive) of feelings	1	[PS03]
Eight emotions (anger, anticipation/expectation, disgust, fear, joy, sadness, surprise, confidence)	3	[PS14], [PS17], [PS18]
Six emotions (anger, disgust, fear, joy, sadness and surprise)	1	[PS02]
Affective states (frustrated, unmotivated, anxious, and bored)	1	[PS09]
only negative emotions	1	[PS11]
They do not make explicit what the emotions would be, they just cite frustration and confusion	1	[PS05]

Source: designed by the authors

4. Discussion

Based on the results found in primary studies, it appears that sentiment analysis in VLE is used to: (1) investigate students' opinions about courses, teachers and teaching materials (PS01, PS08, PS09, PS12, PS13, PS15, PS20); (2) improve the teaching process (PS03, PS11, PS16, PS18); (3) provide personalized learning provide personalized learning (PS06, PSEP10, PS19); (4) improve student performance (PS04, PS05); (5) evaluate the effectiveness and quality of the VLE (PS02, PS07); (6) predict student dropout from the course (PS14); and (7) form work teams (PS17).

In order to gain an in-depth view of sentiment analysis during the teaching and learning process, the focus of this systematic mapping consisted of VLE. Thus, it is possible to identify gaps that can be explored in the future in research on EDM and sentiment analysis in these environments. Some of the most important open questions in different aspects of sentiment analysis are presented below.

From the primary studies found in this systematic mapping, a trend towards machine learning algorithms, such as SVM, Naive Bayes and KNN, which identify emotions and feelings based on student interactions, mainly through forums, is perceived. Such predominance is due to the lower degree of difficulty in capturing the text format. However, for mechanisms in VLE to be able to analyze emotions and feelings in educational scenarios of reality, it is necessary for future research to focus on data produced in real-world teaching and learning environments. Such data concern the behavior and interactions of students at different times when using the VLE, whether synchronous or asynchronous, such as: in group activities, answering an assessment or trying to solve content questions.

The primary studies found do not address the granularity of students' representation of feelings and/or emotions in VLE. If the environment allows the teacher to define the frequency of feedback on students' feelings in weeks, daily or even in hours or minutes, it is possible to constantly monitor the process to discover any deficiencies in teaching and learning.

The predominance of primary studies that explore only the polarity of feelings (positive, negative and neutral) is pointed out. However, few researchers address emotions such as anger, disgust, fear, joy, sadness and surprise. Such emotions can manifest details about aspects that interfere in the cognitive and social development of students. In this sense, research should focus more on the analysis of emotions in VLE and, consequently, improve performance in verifying possible obstacles during the learning process or in discovering the relationship between a student's performance and his/her interest in the course.

It is noteworthy that, by exploring other sources of emotions, such as facial expressions, it is feasible to carry out an improved analysis of students' feelings. In this way, it is possible to understand how affective relationships are built in the learning environment and define which pedagogical strategies the teacher could put into practice. In addition, it is important to verify the feelings and emotions of teachers in VLE as well, which can be significant for the management of the educational institution.

5. Conclusion

This article presents a systematic mapping of the literature on EDM and Sentiment Analysis in VLE. To this end, the process of collecting, evaluating and systematizing primary studies was reported through a broader and more quantitative approach, providing data to identify trends in studies, as well as gaps to be explored (PETERSEN et al., 2015). Thus, it is possible to provide support for future research focused on EDM, in addition to sentiment analysis and VLE.

Regarding the results, quantitative data were evidenced, including the number of studies by selection stages for each scientific literature database and year. Then, the primary studies were synthesized and some characteristics of the works were discussed, according to the research questions. Thus, it was possible to obtain an overview of the current scientific literature on EDM, sentiment analysis and VLE. It is emphasized that the search was reduced to a limited number of searchers and studies on sentiment analysis that were not applied in VLE or that did not contribute with any significant method in the context of this systematic literature mapping were excluded.

It is understood that it is necessary to explore other sources of information to capture students'-feelings and emotions, in addition to forums, as discussed in the primary studies found. As next steps, we intend to carry out a systematic mapping on reports of experiences of analysis of feelings and emotions in VLE that emphasize the pandemic period caused by COVID-19.

It is also important to highlight that social networks linked to the learning environment, through the crossing of information, can enable more sophisticated analyzes of students' feelings and closer to reality. In addition, for the continuity of research in the study area addressed in this systematic mapping, it is suggested to carry out empirical studies comparing the performance of various techniques, algorithms and sentiment analysis libraries applied in VLE. It is hoped that the results disclosed in this article will contribute to future research on Distance Education, especially in investigations with an emphasis on the effects of emotions and feelings on performance in activities and on completion of courses. In addition, the study allows us to discuss the recommendation of appropriate strategies with the socio-affective profile of students in VLE and to act in advance by detecting behavioral signs indicative of dropping out of the course..

References

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.

PURAIVAN, E. *et al.*Emotion-based decision support tool for learning processes: an application with undergraduate 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.

Complementary Reading

[PS01] 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.

[PS02] 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.

[PS03] SPATIOTIS, N. *et al*. Evaluation of an educational training platform using text mining. In: **Hellenic Conference on Artificial Intelligence**. 2018. p. 1-5.

[PS04] KECHAOU, Z.; AMMAR, M. B.; ALIMI, A. M. Improving e-learning with sentiment analysis of users' opinions. In: **EDUCON IEEE Global Engineering Education Conference**. 2011. p. 1032-1038.

[PS05] 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.

[PS06] FEI, H.; LI, H. The study of learners' emotional analysis based on MOOC. In: **International Conference on Cognitive Computing**. 2018. p. 170-178.

[PS07] 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.

[PS08] 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.

[PS09] SELMI, M.; HAGE, H.; AÏMEUR, E. Opinion Mining for predicting peer affective feedback helpfulness. In: **KMIS**. 2014. p. 419-425.

[PS10] 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.

[PS11] 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.

[PS12] 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.

[PS13] 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.

[PS14] 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.

[PS15] 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.

[PS16] 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.

[PS17] 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.

[PS18] 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.

[PS19] CLARIZIA, F. et al. E-learning and sentiment analysis: a case study. In: **International Conference on Information and Education Technology**. 2018. p. 111-118.

[PS20] 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.