

The Role of Polarity and Topic in Private Messages in the Analysis of Academic Trajectories: A Case Study

Gabriella Carvalho¹, Jairo F. de Souza^{1,2}, Renê R. Veloso³

¹ LApIC Research Group – Federal University of Juiz de Fora (UFJF)
Juiz de Fora – MG – Brazil

²Department of Computer Science – Federal University of Juiz de Fora (UFJF)
Juiz de Fora – MG – Brazil

³Department of Computer Science, Center for Exact and Technological Sciences
State University of Montes Claros (Unimontes) – MG – Brazil

carvalho.gabriella@estudante.ufjf.br ,

jairo.souza@ufjf.br ,

rene.veloso@unimontes.br

Abstract. *This case study investigates Moodle private messages as a complementary source for monitoring academic trajectories in distance education courses. Messages from students classified as Graduated and Dropout were analyzed using a combination of sentiment analysis and topic modeling. The results indicate low consensus among the analysis methods and weak correlations between average polarity and indicators. Topic modeling showed that the messages primarily address academic and administrative demands, such as grades, deadlines, and technical difficulties. We conclude that average polarity has limited scope, and that using VLE textual records in real contexts depends on the centrality, quality, and completeness of the recorded interactions.*

1. Introduction

Historically, student behavior has been identified through face-to-face interactions in classrooms, where teachers and administrators can observe signs of disengagement or difficulties faced by students [Mehall 2020]. With the emergence of virtual environments, however, teachers have reported greater difficulty making such observations and monitoring student progress, particularly due to the absence of nonverbal and behavioral cues typical of face-to-face teaching [van Leeuwen 2023].

As education moved into digital environments, this process became known as electronic learning (e-learning) [Zhou et al. 2023]. This transition strengthened the role of Virtual Learning Environments (VLEs), which enabled new forms of interaction between students and teachers [Mora et al. 2017], although these interactions are more mediated and less direct than those occurring in face-to-face education [Olivé et al. 2019].

In remote education, a significant portion of communication among actors, such as students, teachers, and tutors, is expected to occur through technology-mediated texts [Anwar et al. 2023]. However, the interaction dynamics among these actors, the communication practices adopted, and the institutional culture regarding the use of the VLE as

a central communication channel have important implications, particularly for the data generated in these environments. These implications concern the quantity, availability, quality, and analytical potential of such data for identifying student behaviors.

Given this scenario and the large volume of data stored in VLEs, it is feasible to adopt analytical strategies that extract knowledge from these records and support understanding of teaching quality [Martin and Ndoeye 2016, Queiroga et al. 2021]. These data include not only structured records, such as access logs, grades, and activity participation, but also textual data derived from messages and communications exchanged within the environment.

In this context, this study contributes to LA by analyzing Moodle private messages as a complementary source for analyzing the academic trajectories of students classified as Dropout and Graduated, using a real dataset comprising nearly 3,000 messages exchanged between 2017 and 2025 in remotely offered courses. To this end, previously constructed indicators that cover different dimensions of the student experience are analyzed alongside characteristics extracted from students' messages, such as sentiment polarity and recurring topics. Comparing sentiment analysis tools reveals low consensus, as polite language often masks academic struggles. Average polarity proves limited for explaining academic trajectories.

2. Related Work

The field of *Learning Analytics* (LA) has become established as an approach focused on the collection, analysis, and interpretation of educational data to understand and support learning processes [Siemens 2013]. In this context, the construction and use of indicators enable the transformation of raw records from learning environments, such as access logs, grades, and participation data, into informative measures of students' trajectories [Schwendimann et al. 2017, Costas-Jauregui et al. 2021, Menezes 2024, Santos 2024]. As shown in a previous study, indicators related to engagement, performance, motivation, student–teacher interaction, and student dropout can be modeled from Moodle data [Carvalho et al. 2025] to support the interpretation of these educational records.

Despite their usefulness, numerical indicators tend to summarize observable behaviors in the environment, but they do not always capture qualitative aspects of the student experience. Difficulties, frustrations, doubts, requests for help, and signs of disengagement may appear more explicitly in textual interactions, such as private messages and forum posts [Cleveland-Innes and Campbell 2012, Yang et al. 2015, Anwar et al. 2023]. Thus, the analysis of educational texts has been explored as a means of complementing the quantitative measures traditionally used in LA.

Several studies have used students' language to understand patterns of behavior, participation, and academic risk. Bobó et al. [Bobó et al. 2019], for example, propose the SASys architecture, applied to Moodle, to identify emotional profiles from messages posted in question-and-answer forums and combine them with interaction data to identify students who may be at risk of dropping out. In a similar vein, Pfitscher et al. [Pfitscher et al. 2023] analyze textual responses from programming students to support student retention efforts, while Pereira et al. [Pereira et al. 2024] investigate messages from school WhatsApp groups to identify signs of bullying.

Studies of forums in MOOCs and distance education (DE) courses have also iden-

tified textual signals that may be associated with educational phenomena such as dropout and disengagement [Wen et al. 2014, Yang et al. 2015, Kagklis et al. 2015]. In the Moodle context, Dake & Bada [Dake and Bada 2023] demonstrate the potential of this type of analysis by exploring sentiments associated with platform use and proposing a classifier to predict future polarities. Table 1 summarizes the main related studies that used texts produced in VLEs and *Massive Open Online Courses* (MOOCs).

Table 1. Summary of textual corpora used in related studies

Study	Environment	Textual corpus analyzed
[Bobó et al. 2019]	VLE (Moodle)	27 student messages from 10 question-and-answer forums in a distance-learning course.
[Kagklis et al. 2015]	DE course forum	198 messages from 34 students and 173 messages from 4 tutors posted in discussion forums.
[Wen et al. 2014]	MOOC (Coursera)	36,590 discussion forum posts from three MOOCs.
[Yang et al. 2015]	MOOC (Coursera)	12,434 posts from 4,281 students in two courses, combined with platform interaction data.
[Cleveland-Innes and Campbell 2020]	VLE	Transcripts of online discussions and open-ended responses from 217 students in 19 courses from two graduate programs.
[Dake and Bada 2023]	VLE (Moodle)	300 open-ended student responses, collected through a form, about their experience with online assessments in Moodle.
[Lundqvist et al. 2020]	MOOC (Future-Learn)	25,000 online posts and 337 feedback statements produced by 264 participants in a programming MOOC.

The summary presented in Table 1 indicates that the analysis of educational texts has been applied in different contexts, especially in forums and open-ended responses. However, some studies rely on small message sets, whereas those with larger corpora tend to focus on public interactions, such as forum posts. This scope may limit the analysis of institutional contexts in which communication between students and instructors does not occur primarily in these spaces. This study differs from previous work by analyzing real Moodle private messages, written in Portuguese, and by discussing their potential and limitations as a complementary source to educational indicators for understanding the trajectories of students classified as *Dropout* and *Graduated*.

From a technological perspective, the literature shows that sentiment analysis is not restricted to a single technique. Approaches range from lexicon-based methods, which assign polarity based on word dictionaries, to machine learning models and *transformer* architectures, which learn classification patterns from textual representations. Table 2 summarizes some of these approaches identified in the literature for sentiment analysis in messages.

Therefore, there is no single predominant technique for sentiment analysis in educational contexts. While lexicon-based approaches are useful in scenarios without labeled corpora and are simpler to apply, machine learning methods and *transformer* models tend

Table 2. Studies and main techniques used in sentiment analysis

Study	Main technique or tool
[Id et al. 2024]	<i>Transformer</i> models, especially mBERT.
[Koufakou 2023]	RoBERTa for polarity extraction.
[Souza and Vieira 2012]	OpLexicon.
[Freitas 2013]	SentiLex.
[Lundqvist et al. 2020]	VADER.
[Bobó et al. 2019]	Frame Semantics (FrameNet Brasil).
[Pereira et al. 2024]	TextBlob and BERT.

to capture message context more effectively, although they may require greater computational resources. This variation across studies highlights that sentiment analysis results depend on the corpus, the communication context, and the technique employed. For this reason, this study compares different approaches to observe how they behave when applied to Moodle private messages.

Given this background, this case study examines the potential and limitations of Moodle private messages as a complementary source to previously modeled educational indicators, considering students classified as *Dropout* and *Graduated*. To this end, a comparative analysis is conducted across different sentiment analysis techniques, including lexicon-based approaches and *transformer* models, to examine how these methods classify messages and the extent to which the extracted signals relate to educational indicators and students' academic status.

3. Materials and Methods

This section presents the methodology adopted in the study, including the characterization of the dataset, the definition of the analysis groups, the description of the educational indicators used, and the selection of private messages from Moodle. The text preprocessing steps and the sentiment analysis techniques employed are also described.

3.1. Dataset

This study was conducted in the context of the State University of Montes Claros (Unimontes), an institution within the Open University of Brazil (UAB) system. In this model, the didactic and pedagogical organization of distance education (DE) programs involves both mandatory activities in the VLE, such as quizzes, assignments, and forums, and in-person assessments.

The analyzed dataset comprises historical data from 20 DE programs offered by Unimontes in 2018, 2020, and 2023, all mediated through Moodle.¹ In total, approximately 4,500 users linked to the environment were identified, of whom 3,763 were students. Among these students, 39.5% had the status *Enrolled*, 29.7% *Graduated*, and 30.8% *Dropout*. The messages in the dataset span September 2017 to February 2025.

For the application of sentiment analysis techniques, the study focused on students whose academic trajectories had already been defined, considering only the *Graduated*

¹The study was approved by the Research Ethics Committee, under registration CAAE/UNIMONTES 94419625.8.0000.5146.

and *Dropout* groups. After excluding students classified as *Enrolled*, the analytical subset comprised 2,278 students, of whom 50.9% were *Dropout* students and 49.1% were *Graduated* students. This classification of academic trajectory was constructed from academic records originally registered at the course-unit level and later consolidated at the individual level.

3.2. Ethical Considerations

This study is part of a research project approved by a Research Ethics Committee, under registration CAAE/UNIMONTES 94419625.8.0000.5146. The study was conducted in accordance with the ethical guidelines applicable to research involving human participants. To preserve participants' privacy, all analyses were conducted using anonymized data, and no personally identifiable information was disclosed. The results are presented in aggregate form, and the examples of private messages were selected to avoid identifying students, instructors, tutors, or courses.

3.3. Educational Indicators

In this study, the educational indicators previously developed in [Carvalho et al. 2025] are used to monitor students in Moodle. These indicators represent different dimensions of the academic experience, as summarized in Table 3.

Table 3. Summary of the educational indicators used in the study

Indicator	Source / numerical metric	General interpretation
Engagement	Number of posts in assessed forums	Represents student participation in assessed activities in the VLE
Motivation	Number of posts in non-assessed forums	Indicates participation in spaces not directly linked to formal assessment
Performance	Composite <i>score</i> based on the grade obtained and the student's relative position in the class	Represents academic achievement by considering both absolute and comparative performance
Student–teacher interaction	Number of interactions between the student and the instructor in forums	Reflects the intensity of academic interaction between the student and the instructor
Cognitive Depth	<i>Score</i> derived from interaction traces in forums, assignments, and quizzes	Seeks to represent the student's level of engagement with VLE activities
Dropout Risk	Signal constructed from low levels in other indicators	Indicates possible vulnerability to dropout or the need for closer monitoring

Operationally, each indicator is represented by a numerical measure calculated from VLE records and an interpretive classification defined by the student's relative position in the observed distribution. In this study, the numerical measures associated with each indicator were aggregated at the student level across all course units, using totals or medians according to the indicator definition. This aggregation provided a summary of the behavior exhibited throughout each student's academic trajectory and an overall measure of their experience across the analyzed course units.

3.4. Private Messages

The selection of the textual corpus began with extracting Moodle private messages that were not automatic notifications and contained text. Since these messages have a system-wide context, it was necessary to cross-reference them with the institutional academic records, using the CPF identifier as a matching key to retain only records associated with students. Cases in which users, although registered as students at some point, later acted as tutors or instructors were also removed, avoiding the inclusion of communications that did not represent the student experience. Thus, the corpus was constructed through successive refinement steps aimed at preserving interactions effectively attributed to students.

Message preprocessing involved duplicate removal and text normalization, including lowercase conversion, HTML entity decoding, and HTML tag removal. Links, email addresses, long numerical sequences, and excessive whitespace were also removed. Finally, very short messages sent by the same user to the same recipient on the same day were merged whenever possible; those that remained excessively short and could not be grouped were removed from the dataset. The final distribution of textual records by group is presented in Table 4.

Table 4. Communication activity by group in private messages

Group	Number of students with messages	Number of private messages	Average number of messages per student
<i>Dropout</i>	121 (20.03%)	386 (12.99%)	3.19
<i>Graduated</i>	483 (79.97%)	2585 (87.01%)	5.35

An unequal distribution of communication activity can be observed between the groups. In the analyzed context, students who completed the program showed greater communication activity on Moodle, at least in terms of the use of private messages.

3.5. Sentiment Analysis Strategies

As discussed in Section 2, the most common strategies for sentiment analysis can generally be grouped into lexicon-based approaches and machine learning-based approaches. Based on this overview, this study explored the lexical methods SentiLex-PT and OpLexicon, as well as the tools VADER and TextBlob, whose results were later compared with one another. Machine learning-based approaches were then incorporated through the XLM-RoBERTa² and mBERT³ models. This combination of different analytical strategies aimed to compare their behavior and examine the content produced by students from multiple perspectives.

For all sentiment analysis methods, the messages underwent the same text-cleaning process described previously, ensuring a standardized input dataset for comparison across techniques. In the specific case of VADER and TextBlob, the messages were also translated into English since these methods were originally developed for that language.

²cardiffnlp/twitter-xlm-roberta-base-sentiment

³cardiffnlp/mbert-base-tweet-sentiment-pt

3.6. BERTopic

In addition to sentiment analysis, this study sought to investigate the main topics present in the messages. To this end, BERTopic was employed as an unsupervised approach that combines contextualized semantic representations, dimensionality reduction techniques, and clustering algorithms, allowing texts to be organized into thematically similar groups. In this study, this approach was used to analyze the predominant content of students' private messages and to verify whether certain topics occur more frequently in one group.

For the BERTopic application, the messages underwent light preprocessing, primarily for text normalization and basic content standardization. Excessively aggressive cleaning was avoided before embedding generation in order to preserve the semantic context of the messages. Next, the MiniLM model⁴ was used, as it can map sentences and paragraphs into a dense 384-dimensional vector space while preserving semantic information relevant to topic identification.

4. Results

This section presents the results obtained using the method described above. First, the polarities identified by each sentiment analysis method are compared. Next, the relationship between these polarities and the educational indicators considered in the study is discussed. Finally, the topics obtained with BERTopic are presented.

4.1. Sentiment Analysis

The comparison among sentiment analysis methods showed that the polarity assigned to Moodle private messages varies considerably depending on the analytical approach used. As shown in Figure 1, lexicon-based methods tend to assign a higher proportion of positive labels, whereas contextual models based on *transformers* produce distributions that are relatively more concentrated in the neutral and negative classes.

This behavior is consistent with the characteristics of each approach. Lexicon-based methods assign greater weight to the polarity of isolated textual units, which may favor positive classifications in messages containing greetings, expressions of gratitude, and other markers of politeness. However, in the analyzed corpus, these polite expressions often appear in messages whose main content concerns academic difficulties, technical problems, requests for grade review, concerns about grades, or the risk of failing a course.

In contrast, *transformer*-based models analyze the utterance in a more contextualized way, incorporating broader semantic relationships. As a result, they tend to classify part of these messages as neutral or negative, especially when the overall content indicates difficulty, dissatisfaction, or the need for institutional intervention.

To illustrate this behavior, Table 5 presents representative examples of divergence and agreement among the methods. The selected messages refer to recurring situations in the corpus, such as course failure, grade disputes, technical problems in the VLE, and dissatisfaction with instructional materials or assessments.

⁴paraphrase-multilingual-MiniLM-L12-v2

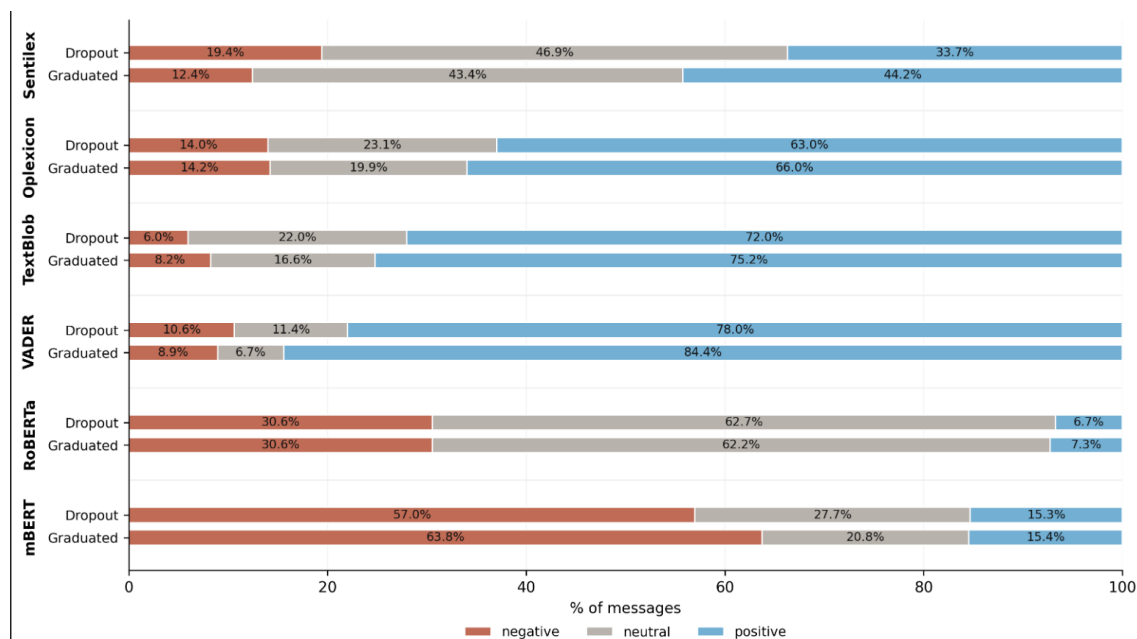


Figure 1. Distribution of sentiment labels by method and group in private messages.

Table 5. Examples of divergence and agreement among sentiment analysis methods. S=Sentilex, O=OpLexicon, T=TextBlob, V=Vader, X=XLM-R, and m=mBert.

Message excerpt	S	O	T	V	X	m	Interpretation
“Bom dia professora, tudo bem? não consegui ser aprovada em sua disciplina, por 0,50 pontos. Terei que refazer a disciplina [...]? aguardo retorno. atenciosamente.” ⁵	+	-	+	+	-	-	The message refers to failing a course by a small grade difference, but includes greetings and a polite closing, which may favor positive classifications in some methods.
“Boa tarde. nota da atividade colaborativa, 0 em 10? sem justificativa, professora? não, ué. assim não tem como prosseguir, por favor me dê um retorno plausível. O porque?” ⁶	+	+	+	+	-	-	Although the message expresses disagreement, dissatisfaction, and a demand for justification, it was classified as positive by lexicon-based methods.
“Não estou conseguindo abrir o fórum [...] aparece erro de codificação.” ⁷	-	+	+	+	-	-	The excerpt describes a technical problem in the VLE, showing that objective messages about platform failures may receive different polarities depending on the method.
“Professora, às vésperas de sua prova eu e a maioria das alunas estamos em dúvida [...] seus vídeos são de péssima qualidade de áudio [...] o conteúdo é muito complicado de entendimento [...] queríamos muito uma webinar.” ⁸	-	-	-	-	-	-	In this case, all methods agree, since the message contains explicit negative markers, such as difficulty, criticism of the material, and dissatisfaction with the assessment.
“Gostaria de te agradecer pelas aulas muito bem explicadas, excelente conteúdo e uma dicção perfeita. Estou realmente encantada pelo seu trabalho. Parabéns de coração!” ⁹	+	+	+	+	+	+	In this case, all methods agree on a positive classification, since the message contains explicit positive markers, such as gratitude, praise, and recognition of the teacher’s work.

⁵ Author translation: “Good morning, professor, how are you? I was not able to pass your course because of 0.50 points. Will I have to retake the course [...]? I look forward to your reply. Sincerely.”

⁶ Author translation: “Good afternoon. Grade for the collaborative activity, 0 out of 10? With no justification, professor? No, come on. It is not possible to continue like this; please give me a plausible answer. Why 0?”

⁷ Author translation: “I am not able to open the forum [...] an encoding error appears.”

⁸ Author translation: “Teacher, on the eve of your exam, most of the students and I are still confused

The examples indicate that divergence tends to occur mainly in formally polite messages, in which institutional language softens situations of academic difficulty or dissatisfaction. Conversely, when the message contains explicit negative markers, such as direct criticism, difficulty understanding the content, or dissatisfaction with materials and assessments, greater convergence among the methods is observed.

Table 6 reinforces this interpretation by showing low consensus among the methods. The percentage of messages with full consensus was low in both groups, corresponding to 4.7% among dropout students and 3.2% among graduated students. In contrast, messages classified into two or three distinct polarity labels predominated, indicating that sentiment analysis in this corpus is sensitive to the method adopted.

Table 6. Level of divergence among sentiment analysis methods by group

Group	Full consensus	Two labels	Three labels
<i>Dropout</i>	18 (4.66%)	174 (45.08%)	194 (50.26%)
<i>Graduated</i>	82 (3.17%)	1007 (38.96%)	1496 (57.87%)

It is also observed that the *Graduated* group presented a higher proportion of messages classified into three distinct polarity labels, suggesting greater instability among the methods. In the *Dropout* group, the percentage of messages with full consensus was slightly higher than that observed among graduated students.

In both groups, however, the low convergence indicates that the affective interpretation of private messages cannot be treated as independent of the chosen analytical approach. Thus, in the analyzed context, message polarity should be interpreted with caution, especially because educational and administrative interactions may combine polite language with concrete situations of academic difficulty, concern, or dissatisfaction.

As a complementary analysis, the effect of polite expressions on lexicon-based methods was investigated. To this end, recurrent expressions used in greetings, gratitude, and formal closings, such as “good morning”, “please”, and “sincerely”, were removed. Of the 2,971 messages analyzed, 2,937 contained at least one of these markers, indicating that politeness is a recurrent feature of Moodle private interactions and is expected in academic and formal contexts. Therefore, in this corpus, these expressions may be treated as domain-specific stopwords.

After this removal, SentiLex and OpLexicon were reapplied, allowing the original polarity to be compared with the polarity obtained without these markers. The results show that 21.91% of the messages changed labels with SentiLex and 42.21% with OpLexicon, indicating that OpLexicon is more sensitive to these terms. The main observed change was from positive to neutral: 15.42% of the messages with SentiLex and 26.99% with OpLexicon. In addition, with OpLexicon, 3.10% of messages shifted from positive to negative. These results suggest that part of the positivity captured by lexicon-based methods was associated with polite communication rather than necessarily with the central academic content of the messages.

[...] your videos have very poor audio quality [...] the content is very difficult to understand [...] we would really like a webinar.”

⁹Author translation: “I would like to thank you for the very well-explained classes, excellent content, and perfect diction. I am truly delighted with your work. Congratulations from the heart!”

4.2. Correlation Between Sentiment Analysis and Educational Indicators

To verify whether the polarities expressed in private messages were related to the previously computed student dimensions, Spearman's correlations were calculated between the sentiment scores and variables related to performance, engagement, motivation, student-teacher interaction, cognitive depth, number of course units, and number of private messages.

Among the 604 students who had exchanged private messages, only 70 (61 *Graduated* and 9 *Dropout*) were retained after applying the full-consensus criterion. This criterion preserved 100 messages, corresponding to 3.36% of the original corpus. Before the correlation analysis, message-level polarity scores were averaged within each student, so that each student contributed a single observation regardless of the number of messages posted. Accordingly, the overall correlations were estimated from 70 student-level observations.

A total of 91 correlations were tested, corresponding to all combinations between 13 indicator-related variables and seven polarity measures. For each association, a two-sided p -value and a 95% confidence interval estimated through bootstrap resampling at the student level were calculated. To account for multiple comparisons, the Bonferroni correction was applied, resulting in an adjusted significance threshold of $\alpha = 0.00055$.

Overall, the observed correlations were weak, with values close to zero across most of the combinations analyzed. The highest coefficient was observed between comparative performance and mBERT ($\rho = 0.273$, 95% CI [0.042, 0.476], $p = 0.022$, $p_{\text{Bonferroni}} = 1.000$). For motivation, the strongest association was observed between the total number of posts and mBERT ($\rho = 0.239$, 95% CI [0.041, 0.426], $p = 0.046$, $p_{\text{Bonferroni}} = 1.000$). Cognitive depth in forums showed a weak negative correlation with SentiLex ($\rho = -0.248$, 95% CI [-0.465, -0.007], $p = 0.039$, $p_{\text{Bonferroni}} = 1.000$).

Although five of the 91 correlations presented unadjusted p -values below 0.05, none remained statistically significant after the Bonferroni correction. Therefore, the results do not provide statistically significant evidence of an association between the average polarity of private messages and the educational indicators analyzed. In addition, the results are affected by the selective and unbalanced sample, particularly by the small number of *Dropout* students retained after the full-consensus criterion.

These results are consistent with the literature, which suggests that the relationship between sentiments expressed in online learning environments and educational indicators is not necessarily direct. Wen et al. [Wen et al. 2014], when analyzing sentiments in student posts at the individual level, observed that the effects were not consistent across the investigated courses, suggesting that, in individualized analyses, the interpretation of polarity depends strongly on the content of the messages, the course context, and the type of interaction analyzed.

In the Moodle context, Bobó et al. [Bobó et al. 2019] identified evidence of emotional profiles and potential dropout risk based on a small number of messages in question-and-answer forums. However, unlike the present study, the authors did not treat polarity in isolation, but combined it with interaction records from the VLE. Thus, while Bobó et al. point to the potential of using textual signals in combination with other educational traces, the results obtained in this study indicate that the average polarity of private

messages, when considered in isolation, does not show a strong association with the educational indicators analyzed.

Therefore, the absence of strong correlations in this study does not imply that private messages are irrelevant for student monitoring. On the contrary, it suggests that average polarity alone may be insufficient to capture complex dimensions of the student experience. Private messages may express specific frustrations that are not necessarily directly reflected in the educational variables analyzed. Thus, the results reinforce the need to interpret sentiment analysis as a complementary layer to educational indicators, especially when combined with topic analysis and with the communication context in which the messages were produced.

4.3. Message Clustering by Topic

This step aimed to identify clusters of semantically similar messages and summarize the most recurrent topics in the private messages. To this end, BERTopic was used. Table 7 presents the main macro-topics identified.

Table 7. Summary and size of the main topics identified by BERTopic

Macro-topic	Topic ID (messages)	Messages, <i>n</i> (%)	Summary
Grades, assessments, and exams	3(162), 12(54), 17(42), 18(40), 21(34), 23(25)	357 (12.0)	Missing grades, in-person assessments, answer keys, recovery exams, and grade review requests.
Forums and collaborative activities	0(323), 1(247)	570 (19.2)	Forum participation, forum grades, posts, and group activities.
Deadlines, submissions, and technical issues	4(144), 5(125), 16(42)	311 (10.5)	Dates, delays, email submissions, posting activities, and internet-related difficulties.
Course content	6(112), 7(103), 9(92), 11(75), 22(26)	408 (13.7)	Content copied from activities or answers submitted by students.
Final paper, internship, and supervision	10(83), 13(51), 14(47)	181 (6.1)	Reports, internships, final papers, supervision, and academic corrections.
Other topics and outliers	2(183), 8(97), 15(42), 19(37), 20(37); -1(748)	1,144 (38.5)	Additional topics not included in the main macro-topics and messages classified as outliers by BERTopic.

The identified macro-topics indicate that private messages were mainly used to address academic and administrative demands, such as grades, deadlines, activity submissions, assessment review requests, technical difficulties, and contact with instructors or tutors. Thus, the corpus seems to reflect concrete situations in students' academic trajectories more than spontaneous affective expressions. Although most messages were from graduated students, dropout students had a higher proportional presence in some topics related to deadlines, submissions, and grades.

Topic clustering helps explain why average polarity alone did not show a strong association with the educational indicators: many messages do not express broad affective states, but rather specific requests related to course organization, assessments, and academic support. This result is close to that of Yang et al. [Yang et al. 2015], who show that more specific textual signals, such as confusion, offer more useful interpretations regarding retention and dropout. Thus, topic modeling complements sentiment analysis by highlighting the types of demands that motivated communication on Moodle and indicating that the interpretation of polarities depends on the content and communicative context of the messages.

5. Conclusion

This study explored Moodle private messages as a complementary source for student monitoring, analyzing them together with previously constructed educational indicators. The results show that average polarity, when considered in isolation, has limited explanatory power. However, textual analysis can contribute to understanding academic and administrative demands, as well as qualitative aspects of students' academic trajectories, especially when interpreted in light of message topics and the context of VLE use.

As a limitation, the analysis considered the average polarity of messages and the aggregated values of the indicators, which may dilute relevant sporadic manifestations throughout the student's trajectory. In addition, the volume of private messages was not evenly distributed across the analyzed groups, since graduated students showed greater communication activity in the environment. It is also important to consider that private messages have an educational and administrative nature, which may make the clear identification of sentiments more difficult. Furthermore, VADER and TextBlob were applied after translating the original Portuguese messages into English. Translation can alter sentiment markers, expressions of politeness, and domain-specific meanings; therefore, the results from these methods should be interpreted with caution.

Furthermore, the results suggest that the potential of textual analysis in LA may be constrained by institutional and organizational practices regarding the use of the VLE. When part of the relevant interactions occurs outside Moodle, such as through WhatsApp, email, or face-to-face support, the textual records available in the environment may represent only a fraction of the student experience. In this sense, the absence of strong correlations does not imply that textual data lack value; rather, it highlights that the analytical value of LA depends on the quality, centrality, and completeness of the records produced within the institutional environment.

Finally, by analyzing a real student dataset, this study highlights both the potential and the limitations of Moodle private messages as a source for student monitoring. These findings may support institutional reflection on the importance of centralizing communication records in the VLE to expand their use in strategies for monitoring, guidance, and student retention.

Acknowledgments

This work was partially supported by the Brazilian agencies CNPq, FAPEMIG, and CAPES.

Declaration on the Use of Artificial Intelligence

We declare that Artificial Intelligence (AI) tools were used to correct grammatical errors.

References

- Anwar, A., Rehman, I. U., Nasralla, M. M., Khattak, S. B. A., and Khilji, N. (2023). Emotions matter: A systematic review and meta-analysis of the detection and classification of students' emotions in stem during online learning. *Education Sciences*, 13(9).
- Bobó, M., Campos, F., Ströele, V., David, J. M., and Braga, R. (2019). Identificação do perfil emocional do aluno através de análise de sentimento: Combatendo a evasão

- escolar. In *Anais do XXX Simpósio Brasileiro de Informática na Educação*, pages 1431–1440, Porto Alegre, RS, Brasil. SBC.
- Carvalho, G., Silva, J. H. N. d. O., Carvalho, Á. M., Barrere, E., et al. (2025). Análise de aprendizagem no moodle: Construção de um painel de dados com indicadores educacionais. In *Simpósio Brasileiro de Informática na Educação (SBIE)*.
- Cleveland-Innes, M. and Campbell, P. (2012). Emotional presence, learning, and the online learning environment. *The International Review of Research in Open and Distributed Learning*, 13:269.
- Costas-Jauregui, V., Oyelere, S., Caussin-Torrez, B., Barros-Gavilanes, G., Agbo, F., Toivonen, T., Motz, R., and Tenesaca, J. (2021). Descriptive analytics dashboard for an inclusive learning environment. In *IEEE Frontiers in Education Conference (FIE)*, pages 1–9.
- Dake, D. and Bada, G. (2023). Unveiling learner emotions: Sentiment analysis of moodle-based online assessments using machine learning. *Journal of Information Technology Education: Innovations in Practice*.
- Freitas, L. (2013). Comparing portuguese opinion lexicons in feature-based sentiment analysis. *International Journal of Computational Linguistics and Applications*, 4:147–158.
- Id, I., Saputra, H., Syamsudhuha, S., Kurniawan, R., and Andriyani, Y. (2024). Sentiment analysis of student evaluation feedback using transformer-based language models. In *Indonesian Journal of Electrical Engineering and Computer Science*, 36:1127.
- Kagklis, V., Karatrantou, A., Tantoula, M., Panagiotakopoulos, C., and Verykios, V. (2015). A learning analytics methodology for detecting sentiment in student fora: A case study in distance education. *European Journal of Open, Distance and E-Learning*, 18.
- Koufakou, A. (2023). Deep learning for opinion mining and topic classification of course reviews. *Education and Information Technologies*, 29(3):2973–2997.
- Lundqvist, K., Liyanagunawardena, T., and Starkey, L. (2020). Evaluation of student feedback within a mooc using sentiment analysis and target groups. *The International Review of Research in Open and Distributed Learning*, 21.
- Martin, F. and Ndoeye, A. (2016). Using learning analytics to assess student learning in online courses. *Journal of University Teaching and Learning Practice*, 13.
- Mehall, S. (2020). Purposeful interpersonal interaction: What is it and how is it measured? *Online Learning*, 24.
- Menezes, M. F. d. (2024). Proposta de um dashboard para análise e visualização de dados educacionais de sistemas tutores inteligentes. *Universidade Federal do Amazonas*.
- Mora, H., Ferrández, A., Gil, D., and Peral, J. (2017). A computational method for enabling teaching-learning process in huge online courses and communities. volume 18.
- Olivé, D. M., Huynh, D. Q., Reynolds, M., Dougiamas, M., and Wiese, D. (2019). A supervised learning framework: using assessment to identify students at risk of dropping out of a mooc. volume 32, pages 9–26. Springer Science+Business Media, LLC, part of Springer Nature.

- Pereira, M. R., Alves, L., and Padilha, T. P. (2024). Investigando o uso de análise de sentimentos para identificar comportamentos de bullying em grupo de whatsapp escolar. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 3059–3065, Porto Alegre, RS, Brasil. SBC.
- Pfitscher, R., Camargo, L., Moreira, B., Wang, C., Zedral, R., and Garcia, T. (2023). Análise de sentimentos em turmas de programação com vistas ao apoio à permanência estudantil. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1329–1340, Porto Alegre, RS, Brasil. SBC.
- Queiroga, E. M., Enríquez, C. R., Cechinel, C., Casas, A. P., Paragarino, V. R., Bencke, L. R., and Ramos, V. F. C. (2021). Using virtual learning environment data for the development of institutional educational policies. *Applied Sciences*, 11(15).
- Santos, T. C. B. d. (2024). Lumilab: um dashboard de learning analytics para apoio a gestores e professores de moocs da plataforma lúmina. In *Programa de Pós-Graduação em Informática na Educação (UFRGS)*.
- Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., and Dillenbourg, P. (2017). Perceiving learning at a glance: a systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, 10(1):30–41.
- Siemens, G. (2013). Learning analytics the emergence of a discipline. *American Behavioral Scientist*, 57:1380–1400.
- Souza, M. and Vieira, R. (2012). Sentiment analysis on twitter data for portuguese language. pages 241–247.
- van Leeuwen, A. (2023). Teachers’ experiences of monitoring their students in online higher education: recommendations for course design and opportunities for learning analytics. *Technology, Pedagogy and Education*, 32(5):589–604.
- Wen, M., Yang, D., and Rosé, C. (2014). Sentiment analysis in mooc discussion forums: What does it tell us?
- Yang, D., Wen, M., Howley, I., Kraut, R., and Rose, C. (2015). Exploring the effect of confusion in discussion forums of massive open online courses. pages 121–130.
- Zhou, Y., Zhao, J., and Zhang, J. (2023). Prediction of learners’ dropout in e-learning based on the unusual behaviors. volume 31, pages 1796–1820. Routledge.