ASCILITE 2025
Future-Focused:
Educating in an Era of Continuous Change

Leveraging student confusion in online forum posts to enhance student engagement using text-based learning analytics

Sisi Liu, Rhoda Abadia, Antonella Strambi, David Caldwell, Xueer Caiwei
University of South Australia

Online discussion forums serve as a natural repository of students' emotions and feelings in online learning environments, as students often post to seek assistance with academic challenges and difficulties. Understanding the complexity of student emotions or affective states is crucial for fostering meaningful engagement and promoting academic success. This study focuses on confusion, a cognitive-affective state that emerges from complex learning processes and is recognised as potentially beneficial for deep learning when reaching an optimal state in terms of its duration and intensity. A text-based learning analytics (TLA) workflow grounded on the confusion and affect dynamics model was proposed. Six machine learning models were evaluated for their effectiveness in classifying confusion from online discussion forum posts collected from three courses. Results indicated that model performance was inconsistent. To demonstrate the viability of the proposed TLA workflow, it was applied to a comprehensive set of course data gathered from two study periods. Through systematic evaluation of students' confusion states and transitions and their relationship with students' academic performance, the study establishes the feasibility of the proposed workflow in facilitating educators' early identification of students experiencing prolonged cognitive disequilibrium that may lead to unproductive struggle rather than meaningful learning gains.

Keywords: confusion, student engagement, text-based learning analytics, online learning

Introduction

The rapidly evolving landscape of online learning environments in higher education has fundamentally transformed how students engage with course materials and interact with peers and educators. An online discussion forum is one of the most frequently used communication tools because of its flexibility and effectiveness in facilitating interactions and collaborations (Nor et al., 2010; Onyema et al., 2019). Students post to seek help with troubles or challenges they are experiencing, and students' posts often contain information about whether and how they are learning and feel supported in their learning journey. Educators can use such information to create the best possible learning environment for their students and provide support tailored to individual students' needs. Similar to traditional classroom learning, understanding the dynamics of student emotions or affective states plays a pivotal role in fostering healthy student engagement and promoting academic success in online learning environments (Arguel et al., 2019; Artino, 2012).

Within the realm of achievement and epistemic emotions, confusion represents a particularly noteworthy type, as research has revealed its evolution from being consistently detrimental to learning to serving as a catalyst for complex learning and a critical cognitive disequilibrium in affect transitions (Arguel et al., 2019; D'Mello & Graesser, 2012; D'Mello et al., 2014; Lodge et al., 2018; Pekrun, 2006). The complexity of confusion becomes particularly pronounced in online learning environments where students lack immediate facial cues and real-time feedback from educators, owing to most of the interactions and communications happening through written text (Arguel et al., 2017; Ma et al., 2024). Limited visual cues and delay in feedback make it challenging to distinguish between productive confusion that promotes deep learning and unproductive confusion that leads to frustration and disengagement (D'Mello et al., 2014). Consequently, the need for appropriate technological scaffolding and timely instructional support is compelling. Although online discussion forums automatically collect student learning data whenever a post is received from a student, manually analysing such data to contextualise student confusion experienced through the course delivery is time- and resource-intensive in courses with large enrolments.

A possible solution is the use of Natural Language Processing (NLP) and machine learning techniques to support learning analytics in detecting confusion in online discussion forum posts. For instance, Yang et al. (2015) utilised text mining techniques to analyse linguistic features in student interactions that correlate with confusion, such as the use of hedging language or an increase in question frequency. Geller et al. (2020) developed a rule-based approach, in which the rules were formulated based on students' self-reported affective states using a set of pre-defined hashtags, that informed the design of an automated classifier for confusion detection in online forms. Later in 2021, Geller et al. (2021) refined their previous work by adopting a BERT-based model for detecting confusion of various types. In their study, they also proposed a labelling tree for confusion to guide the manual annotation of confusion in student posts. However, due to their generality in methods and limited reference to relevant educational theories, these studies lack conceptual grounding to explain how confusion detected affect the underlying learning mechanisms, and miss opportunities to connect their findings to broader pedagogical principles, e.g., constructive confusion resolution (D'Mello & Graesser, 2012) and productive struggle scaffolding (Ma et al., 2024), that could inform more systematic and transferable approaches to educational practice.

To address these limitations, this study proposes a text-based learning analytics (TLA) workflow, grounded on the confusion in learning and productive struggle model (D'Mello & Graesser, 2012), that leverages confusion in text generated by students during online interactions to facilitate educators' early identification of students who may be struggling with a task, and to offer appropriate and just-in-time scaffolding to help students overcome the challenge and to promote more healthy engagement with the course materials. Forum post data was collected and manually annotated to evaluate the performance of several existing models based on the accuracy of the confusion classification results. The best performing model was then incorporated into the proposed TLA workflow to further assess its feasibility on a larger-scale course data as a proof of concept.

Theoretical framework

This study is theoretically grounded on the confusion and affect dynamics model introduced by D'Mello et al. (2014), which represents a significant framework for understanding how the transitions of cognitive and affective states impact student learning. In contrast to assumptions about confusion being a barrier to learning, the model identifies confusion as a natural cognitive-affective disequilibrium that emerges when learners encounter information that contradicts their existing knowledge, contains inconsistencies, or presents concepts that don't easily fit into their current mental frameworks. The authors state that when confusion is optimal, it makes a positive contribution to learning and is one of the emotions most often experienced by students. They also suggest that confusion is necessary for complex learning, and students who experience confusion and are then able to resolve it learn more or even better than those who do not experience this (D'Mello et al., 2014).

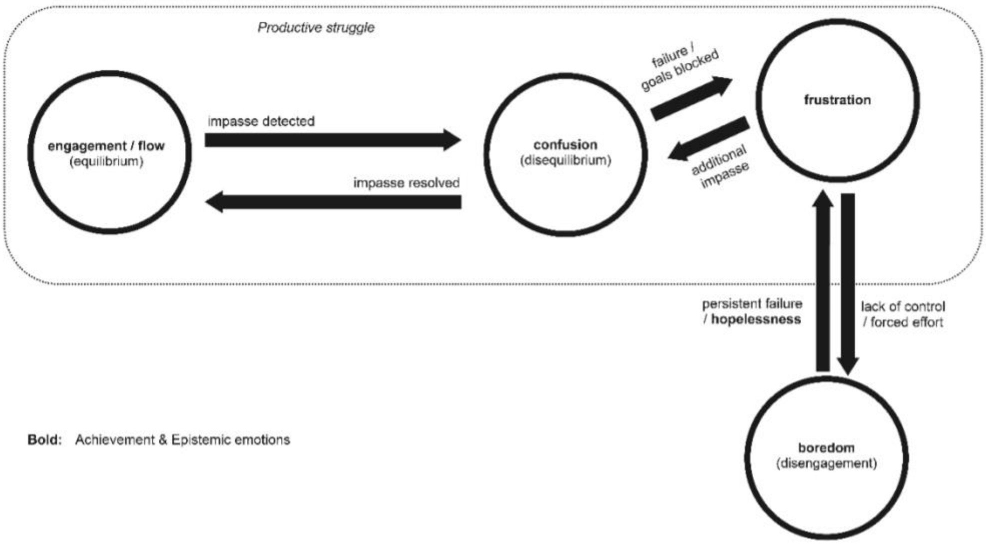


Figure 1. Confusion and the productive struggle model. Image adapted from D'Mello and Graesser (2012)'s affect dynamics model diagram with modification

In their "affect dynamics model" diagram (see Figure 1), D'Mello et al. (2014) argued that productive confusion occurs when learners experience an optimal level of cognitive disequilibrium. The process to determine productive confusion and unproductive confusion generally unfolds in four stages: (a) Students encounter materials that don't align with their expectations or prior knowledge; (b) Their mind recognise inconsistencies and begins actively working to resolve them; (c) Students engage in various strategies to make sense of the challenging or conflicting information; (d) They either resolve confusion into understanding (productive) or confusion persists and becomes frustration or boredom (unproductive). As the process ties closely to the duration and intensity of confusion, the temporal dynamics become crucial in determining whether confusion is enough to challenge students' existing understanding but not so much as to cause overwhelming frustration or disengagement.

Method

In this study, a TLA workflow integrated with NLP and machine learning techniques was developed to identify students experiencing confusion and explore the confusion dynamics in online discussion forum posts. The analytical results can be utilised by educators to flag unresolved or prolonged confusion and offer timely intervention and scaffolding to students to keep them engaged.

Data

A collection of course online forum post data was obtained from three fully asynchronous online programming courses, denoted as **Course 1**, **Course 2**, and **Course 3,** respectively, offered by the University of South Australia Online (UO). For the evaluation of model performance of confusion classification, sample data from the data collection were manually annotated by two linguistic experts as either 'Confusion' or 'Non-confusion'. The process was done iteratively until reaching a 100% inter-rater agreement. A summary of the data labels for each course is described in Table 1.

Table 1
Summary of the data labels

	Confusion	Non-confusion	Total
Course 1	49	52	101
Course 2	136	69	205
Course 3	96	51	147

Classification models

As mentioned earlier, manually analysing online discussion forum data can be time- and resource-intensive, particularly in courses with large enrolments. Therefore, machine learning classification models were first examined and incorporated into the workflow to allow automated determination of whether a post expresses confusion or not.

Utilising NLP techniques, annotated data was transformed into numeric representations that can be recognisable and processed by a classification model. A model then distinguishes between confusion and non-confusion by analysing language patterns derived from the numerically represented data.

Six classification models were selected, with SVM (Cortes & Vapnik, 1995) and LSTM (Graves & Graves, 2012) chosen as the baseline models, and four BERT-based models, including BERT (Devlin et al., 2019), DistilBERT (Sanh et al., 2019), EduBERT (Clavié & Gal, 2019), EduDistilBERT (Clavié & Gal, 2019), chosen for their reported effectiveness in text classification tasks (Clavié & Gal, 2019). SVM is a traditional machine learning model that's computationally efficient and works well with smaller datasets, but its performance is highly subjective to the quality of feature engineering and generally cannot handle complex textual relationships well. LSTM is a recurrent neural network model that can capture sequential patterns in text, but it requires more computational resources than SVM and performance-wise less effective than the BERT-based models. The four

BERT-based models offer state-of-the-art performance on most NLP tasks through the implementation of attention mechanism for better understanding of the bidirectional context. However, BERT-based models require significant computational resources and memory for training and inference. In comparison to BERT, DistilBERT is more balanced regarding its performance and complexity. Meanwhile, EduBERT and EduDistilBERT are preferred options due to being pre-trained on education content that provides the models with domain-specific knowledge and more likely to yield superior performance in educational contexts.

Model hyperparameters, the configuration options for training a model, were examined for a few iterations until the best results were yielded. Table 2 lists the final selection of the hyperparameters for all models, excluding SVM, for which a grid search approach was used. Specifically,

- Max_length/Hidden_size, in which max_length represents the maximum sequence length for a post to be used as input data where a post longer than the maximum sequence will be truncated. Hidden_size represents the size of hidden layers in the transformation architecture (only required for the LSTM model). Values were shared for max_length and hidden_size.
- Epoch represents the number of iterations through the entire set of posts used for training.
- Batch_size represents the number of sample posts processed in each training iteration.
- Learning_rate represents the step size for parameter updates during optimisation. It determines whether an optimal result can be achieved.
- Weight_decay represents the degree a model preserves its previous certainties with each update. It prevents a model from becoming too certain about complex patterns, thus affecting the model's generalisation ability.
- Optimiser is a study method used to determine how a model learns about complex patterns.

Table 2
Hyperparameters for classification models

	Max_length/ Hidden_size	Epoch	Batch_size	Learning_rate	Weight_decay	Optimiser
SVM	-	-	-	-	-	-
LSTM	300	12	8	5e-5	-	Adam
BERT	300	6	8	5e-5	0.01	AdamW
DistilBERT	300	6	8	1e-4	0.01	AdamW
EduBERT	300	6	8	5e-5	0.01	AdamW
EduDistilBERT	300	18	8	1e-5	0.05	AdamW

Due to the sample size and imbalanced label distribution, experiments were conducted using 5-fold cross-validation, with each model trained and validated on 80% of the data and tested on 20% of the data. A random selector has been fixed on a seed value of 42 to ensure all models were tested on the same data.

Proposed TLA workflow

Theoretically informed by the confusion and productive struggle model, a TLA workflow is proposed to classify confusion in student-generated text and incorporate a time component to suggest the duration of the confusion. Details of the workflow are presented in Figure 2. In brief, there are three main phases in the process: (1) The data preprocessing phase involves the transferring of data into a tabular format for easier manipulation and removal of all personal identifiable information to ensure data privacy and integrity. (2) The post order labelling phase involves the identification of the number of posts by a user under a subject and their sequences of posting. If a user only posts once under a subject, then the post will be labelled as **Single post**. Alternatively, if a user posts more than once under a subject, then the first post created will be labelled as **First post,** and the rest will be labelled as **Subsequent post**. (3) The confusion classification phase involves the transformation of data into vectors for classification into either **Confusion** or **Non-confusion**.

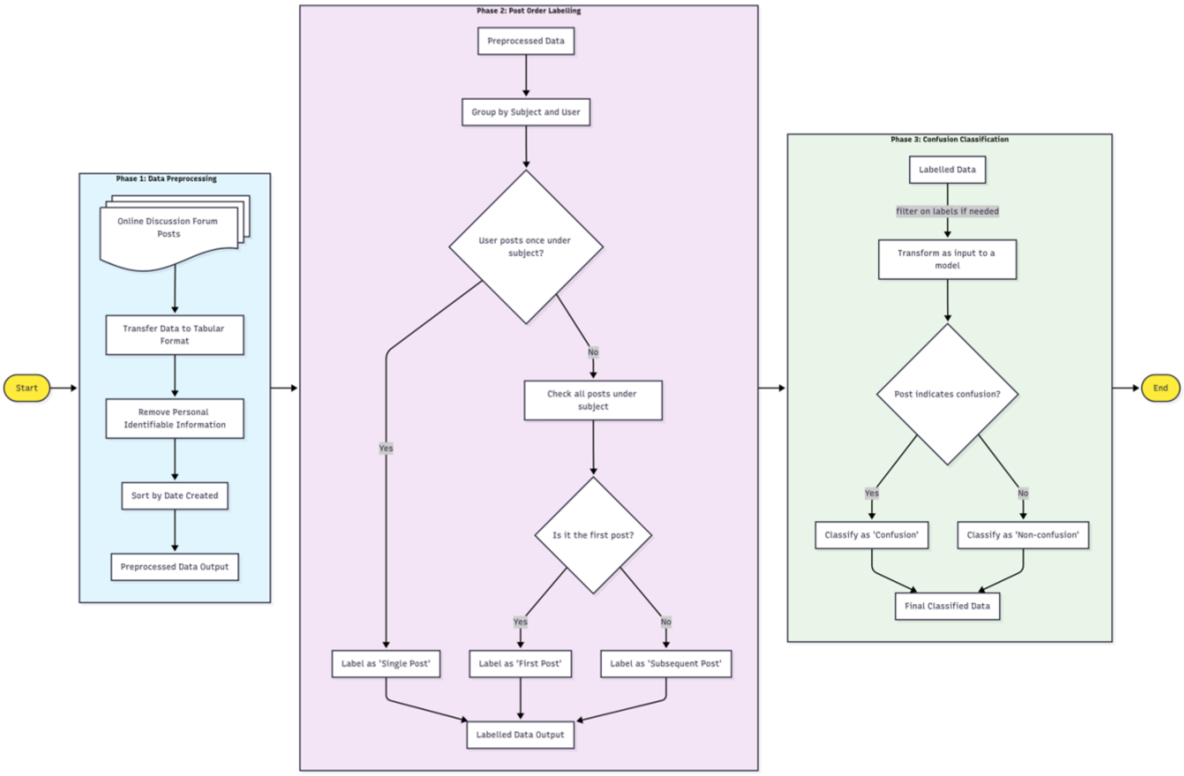


Figure 2. A flowchart diagram of the proposed text-based learning analytics process to leverage student confusion in online discussion forum posts

Note that before transforming the labelled data (after Phase 2) as input to a classification model, an optional filtering step is included to allow flexible adoption of the workflow in courses having a small, medium, or large number of students and in different milestones during a course delivery. For example, when adopting the workflow in a course with an average of hundreds of posts per week around the mid-week in a study period, keeping the subsequent post group only for confusion classification allows more efficient analysis of the posts and effective distribution of the teaching resources to students in need.

Results and discussions

Confusion classification

Model performance was evaluated using the standard metrics, including precision, recall and F1 values, considering the imbalanced label distribution of the data. Specifically,

- Precision represents the percentage of posts correctly classified as confusion out of all posts classified as confusion. Mathematically, $Precision = \frac{Correctly\ classified\ confusion}{All\ classified\ confusion} * 100$
- Recall represents the percentage of posts correctly classified as confusion out of all posts annotated as confusion. $Recall = \frac{Correctly\ classified\ confusion}{All\ annotated\ confusion} * 100$
- F1 represents the weighted harmonic mean of precision and recall for confusion and non-confusion. $F1 = [\frac{\left[\left(2*\frac{Precision(confusion)}{Recall\ (confusion)}*Annotated\ confusion\right) + \left(2*\frac{Precision(non-confusion)}{Recall(non-confusion)}*Annotated\ non-confusion\right)\right]}{Total\ number\ of\ posts}] * 100$

As shown in Table 3, none of the selected models obtained the best performance on all three courses. Individually, BERT performed the best on Course 1 with the highest values in all three metrics, i.e., a precision value of 92.06%, a recall value of 90.48%, and an F1 value of 90.43%. EduBERT performed the best on Course 2 with the highest value in two metrics, i.e., a precision value of 92.41% and an F1 value of 90.46%, whereas both BERT and EduBERT obtained a recall value of 90.24%. As for Course 3, although EduDistilBERT performed slightly better with the highest F1 value of 80.38%, both EduBERT and EduDistilBERT obtained a recall value of 80% and DistilBERT obtained the highest precision value of 82.22%.

Table 3
Performance metrics for confusion classification (best performing results were highlighted in bold)

	Course 1			Course 2			Course 3		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
SVM	67.35%	67.33%	67.33%	48.97%	49.27%	48.82%	45.76%	43.02%	45.61%
LSTM	83.63%	76.19%	74.43%	43.37%	65.85%	52.30%	79.31%	70.00%	60.48%
BERT	**92.06%**	**90.48%**	**90.43%**	91.50%	**90.24%**	89.77%	72.85%	70.00%	70.74%
DistilBERT	84.13%	76.19%	75.07%	87.69%	87.80%	87.69%	**82.22%**	76.67%	77.33%
EduBERT	86.10%	85.71%	85.71%	**92.41%**	**90.24%**	**90.46%**	79.55%	**80.00%**	79.37%
EduDistilBERT	86.39%	80.95%	80.42%	80.42%	80.49%	79.53%	81.48%	**80.00%**	**80.38%**

The above results prove the effectiveness of BERT-based models in classifying confusion from text compared to the baseline SVM and LSTM models. However, the inconsistency of the model performance implies the difficulty of model generalisation when applying to educational data, which further suggests the need for large volumes of high-quality training data and the development of more robust and adaptive models.

Implementation of the TLA workflow

The proposed TLA workflow was evaluated on a set of 1,078 student posts collected from Course 1 offered in two study periods - Study Period 3 and Study Period 6, denoted as **SP3** and **SP6**. Student grade information (denoted as **HD, D, C, P1, P2, F1, F2**) was appended to support the proof of concept by connecting the dynamics of confusion with academic performance. Following the confusion classification results (see Table 3), a BERT model with the same training configurations (as described in Table 2) and weights obtained from the best-performing model was used in Phase 3.

Under the assumptions that all posts under a subject discuss the similar topics and all users posted to similar subjects are treated individually, among the three post groups, i.e., single, first and subsequent post, labelled in Phase 2, first and subsequent post groups were chosen as the focus of the following analyses given that the relationship between the first and subsequent posts better reflects the duration of confusion expressed by a single user and supports the justification of the confusion and productive struggle model (see Figure 1).

The validity of the confusion classification results was assessed first. Table 4 reports the model classification confidence grouped into higher than and equal to 95% (denoted as **>=95%**) and lower than 95% (denoted as **<95%**).

Table 4
Summary of the percentages of posts based on the classification confidence

	SP3		SP6	
	Confidence(>=95%)	Confidence(< 95%)	Confidence(>=95%)	Confidence(< 95%)
Single post	93.22%	6.78%	83.59%	16.41%
First post	94.03%	5.97%	91.95%	8.05%
Subsequent post	90.72%	9.28%	76.67%	23.33%

In general, the classification results reveal a higher confidence level for the first post group as compared to the subsequent post group. This observation indicates that when a subsequent post exists, the student's first post tends to have a clear indication of confusion or non-confusion. Meanwhile, the subsequent posts can be in more blurred states, e.g., confusion, resolution, or a mix, adding the challenges in facilitating an optimal level of cognitive disequilibrium and developing a strategic resolution mechanism.

To delve deeper into the confusion states and transitions from the first post group to the subsequent post group and its influence on student academic performance, a 100% stacked column chart of the grade distribution of subsequent posts classified as confusion was presented in Figure 3 and a timeline of the post created in the first and subsequent post groups for each grade was shown in Figure 4.

As indicated in Figure 3, the grade distribution was discrepant across two study periods, with a relatively high percentage of F1 (42.86%), i.e., failing students, in SP3 and a high percentage of HD (60.14%), i.e., high-performing students, in SP6. There was no significant pattern, except that students with borderline grades (P1 and P2) consistently had minimum contributions in the online discussion forums, as observed regarding the overall grade distribution in the subsequent post group, where posts were classified as Confusion.

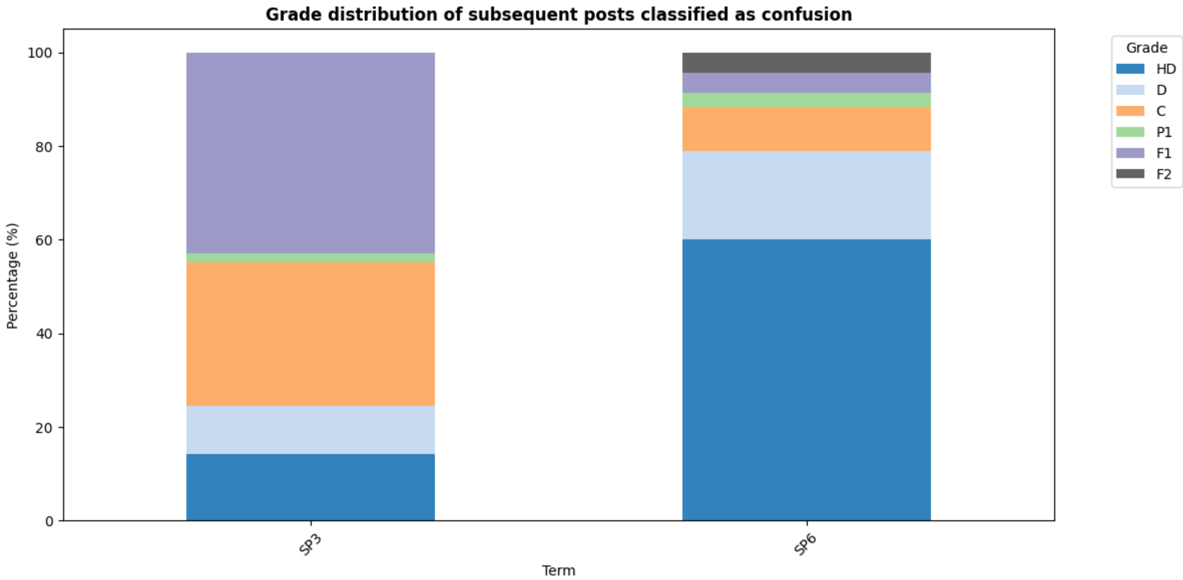
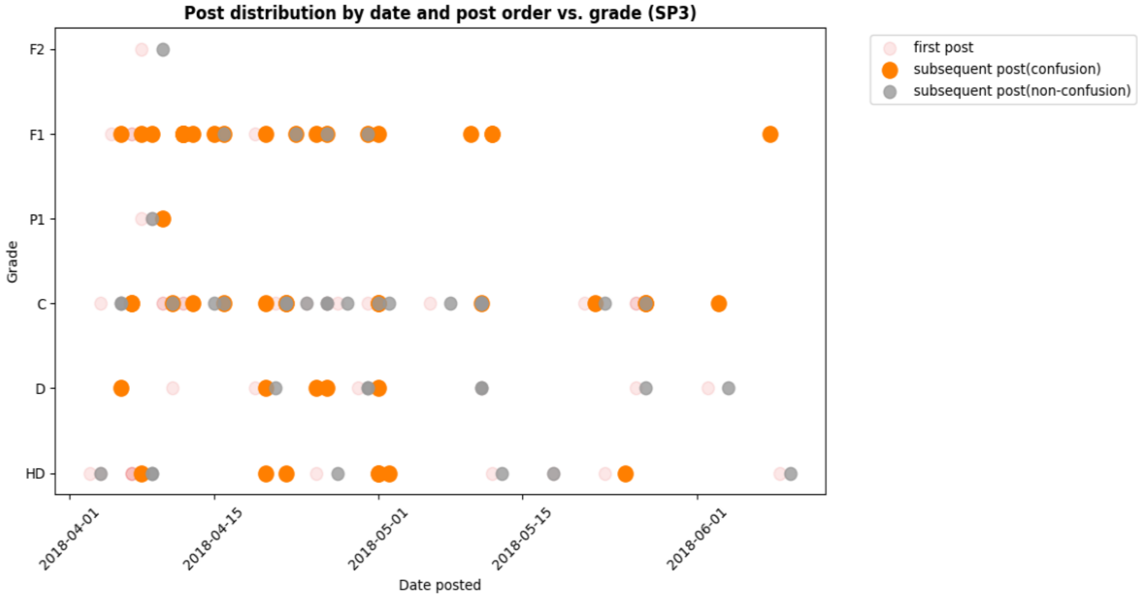


Figure 3. Grade distribution as in percentages of subsequent posts classified as confusion for each term

Figure 4 presents the final set of analysis where a time component was included to provide an in-depth investigation into the confusion transitions in relation to the first and subsequent post groups. Overall, more sparse distributions of the posts were observed among students having lower grade bands, i.e., P1, F1 and F2, indicating less active engagement as compared to the high-performing students. Also, there was a tendency of disengagement towards the latter period of course delivery. Additionally, confusion was comparatively prolonged among these students as shown by the wider gaps among the first posts (marked in translucent orange color) and the subsequent posts classified as confusion (marked in solid orange color). Although similar situations existed in C and HD grade students, they were observed more in the early period of the course delivery.



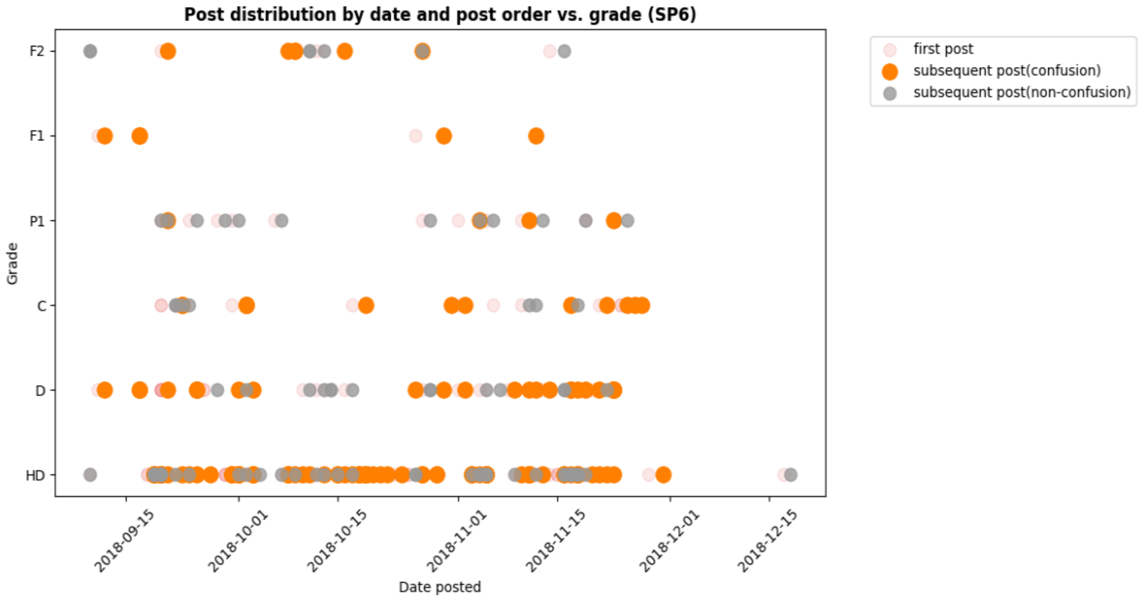


Figure 4. A timeline of the posts created grouped by post order and grade. In the figure, only the first and subsequent posts were included where subsequent posts classified as confusion were highlighted in solid orange color with a bigger marker size

These observations are not surprising as it is well-known that students who engage less with online resources tend to be more at risk of failure or attrition. And the results are well-aligned with the aim of the study where being able to identify persisting confusion in students' posts may assist educators in early identification of at-risk students and enable them to intervene and offer scaffolding to prevent student disengagement.

Conclusions

In summary, through proposing a theory-informed TLA workflow that leverages confusion in text generated by students during online interactions, this study underscores the critical need for the early identification of students who may be struggling and the provision of timely scaffolding to aid their learning journey in online environments. This integration of theoretical foundations with TLA represents a significant advancement in the research on learning analytics, moving beyond purely technical solutions to create meaningful educational interventions grounded in sound pedagogical theory. By bridging the gap between computational analysis and educational practice, this work contributes to the growing body of evidence supporting the integration of theoretical frameworks in learning analytics implementations.

As shown from the confusion classification results, all best-performing models were capable of correctly classifying above 80% of the posts across three courses, which implies that machine learning models are effective in automated confusion detection when integrated with TLA for practical implementation in educational settings. Results also indicate that model generalisation remains a challenge, which can be improved by the utilisation of domain adaptation techniques that account for varying linguistic patterns and cultural differences in student expression, and the development of robust feature extraction methods that can identify universal markers of confusion across different learning environments and student populations.

Moreover, this study reinforces the essential nature of evidence-based learning support and interventions that can transform potentially detrimental cognitive-affective states into productive learning opportunities. Rather than viewing confusion as merely an obstacle to overcome, the findings from applying the proposed TLA workflow to real course delivery data support the need to develop a pedagogical approach of leveraging confusion as a catalyst for deeper learning. This perspective aligns with the contemporary understanding of productive struggle and the beneficial role of cognitive conflict in learning processes.

In future, a user-friendly graphical user interface (GUI) is to be developed to facilitate broader adoption of this workflow and support the visual interpretation of the results, enabling educators to easily implement the proposed system and establish effective warning mechanisms for timely intervention. With the support of the GUI, an action research methodology can be utilised to further investigate the proposed TLA workflow on a larger scale and derive valuable insights into its effectiveness across diverse learning contexts.

References

Arguel, A., Lockyer, L., Kennedy, G., Lodge, J. M., & Pachman, M. (2019). Seeking optimal confusion: a review on epistemic emotion management in interactive digital learning environments. *Interactive Learning Environments*, *27*(2), 200-210. https://doi.org/10.1080/10494820.2018.1457544

Arguel, A., Lockyer, L., Lipp, O. V., Lodge, J. M., & Kennedy, G. (2017). Inside Out:Detecting Learners' Confusion to Improve Interactive Digital Learning Environments. *Journal of Educational Computing Research*, *55*(4), 526-551. https://doi.org/10.1177/0735633116674732

Artino, A. R. (2012). Emotions in online learning environments: Introduction to the special issue. *The Internet and Higher Education*, *15*(3), 137-140. https://doi.org/https://doi.org/10.1016/j.iheduc.2012.04.001

Clavié, B., & Gal, K. (2019). EduBERT: Pretrained Deep Language Models for Learning Analytics. *ArXiv*, *abs/1912.00690*.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*, 273-297.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145-157. https://doi.org/https://doi.org/10.1016/j.learninstruc.2011.10.001

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153-170. https://doi.org/https://doi.org/10.1016/j.learninstruc.2012.05.003

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers),

Geller, S. A., Gal, K., Segal, A., Sripathi, K., Kim, H. G., Facciotti, M. T., Igo, M., Hoernle, N., & Karger, D. (2021). New Methods for Confusion Detection in Course Forums: Student, Teacher, and Machine. *IEEE Transactions on Learning Technologies*, *14*(5), 665-679. https://doi.org/10.1109/TLT.2021.3123266

Geller, S. A., Hoernle, N., Gal, K., Segal, A., Zhang, A. X., Karger, D., Facciotti, M. T., & Igo, M. (2020). *#Confused and beyond: detecting confusion in course forums using students' hashtags* Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, Frankfurt, Germany. https://doi.org/10.1145/3375462.3375485

Graves, A., & Graves, A. (2012). Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, 37-45.

Lodge, J. M., Kennedy, G., Lockyer, L., Arguel, A., & Pachman, M. (2018). Understanding Difficulties and Resulting Confusion in Learning: An Integrative Review [Review]. *Frontiers in Education*, *Volume 3 - 2018*. https://doi.org/10.3389/feduc.2018.00049

Ma, Y., Song, Y., Celepkolu, M., Boyer, K. E., Wiebe, E., Lynch, C. F., & Israel, M. (2024). Automatically detecting confusion and conflict during collaborative learning using linguistic, prosodic, and facial cues. *arXiv preprint arXiv:2401.15201*.

Nor, N. F. M., Razak, N. A., & Aziz, J. (2010). E-learning: Analysis of online discussion forums in promoting knowledge construction through collaborative learning. *WSEAS Transactions on Communications*, *9*(1), 53-62.

Onyema, E. M., Deborah, E. C., Alsayed, A. O., Noorulhasan, Q., & Sanober, S. (2019). Online discussion forum as a tool for interactive learning and communication. *International Journal of Recent Technology and Engineering*, *8*(4), 4852-4859.

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315-341. https://doi.org/10.1007/s10648-006-9029-9

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Yang, D., Wen, M., Howley, I., Kraut, R., & Rose, C. (2015). *Exploring the Effect of Confusion in Discussion Forums of Massive Open Online Courses* Proceedings of the Second (2015) ACM Conference on Learning @ Scale, Vancouver, BC, Canada. https://doi.org/10.1145/2724660.2724677

Liu, S., Abadia, R., Strambi, A., Caldwell, D. & Caiwei, X. (2025). Leveraging student confusion in online forum posts to enhance student engagement using text-based learning analytics. In S. Barker, S. Kelly, R. McInnes & S. Dinmore (Eds.), *Future-focused: Educating in an era of continuous change*. Proceedings ASCILITE 2025. Adelaide (pp. 221-230). https://doi.org/10.65106/apubs.2025.2642