Check for
updates

Automatic content analysis of asynchronous discussion forum transcripts: A systematic literature review

Mubarik Ahmad[1] · Kasiyah Junus[1] · Harry Budi Santoso[1]

Abstract

In recent years, the use of asynchronous discussion forums in online learning has increased rapidly. In earlier studies, content analysis is the most-used research method in exploring the discussion transcripts. However, conventional content analysis is a time-consuming task and requires experienced coders. There is a need for an automated approach to analyse the online discussion transcripts to help instructors optimise the learners' learning experiences. This article presents a systematic literature review of the automated content analysis approach in online discussion transcripts. Fifty-four relevant peer-reviewed conference and journal papers were found between January 2016 and October 2021, using the PRISMA and snowball methods. Eight measurement dimensions were studied from online learning transcripts: cognitive, social, relevance/importance, summary, pattern, behaviour, topic, and learning resources. Six theoretical frameworks were used in the selected studies, namely the Community of Inquiry, Stump's Post Classification, Arguello's Speech Acts, ICAP, Wise's Speaking Behaviour, and Bloom's Taxonomy. All selected studies are experimental, with 93% using machine learning to analyse discussion transcripts. English is the most-used language dataset, used in 78% of studies. These studies reported promising results in accuracy and precision. However, this research area still has room for improvement, especially in the reliability and generalisability of cross-domain context.

✉ Kasiyah Junus
kasiyah@cs.ui.ac.id

Mubarik Ahmad
mubarik.ahmad91@ui.ac.id

Harry Budi Santoso
harrybs@cs.ui.ac.id

[1] Digital Library and Distance Learning Laboratory, Faculty of Computer Science, Universitas Indonesia, Depok Campus, Depok, Jawa Barat, Indonesia 16424

⚘ Springer

1 Introduction

Collaboration is becoming increasingly vital nowadays, especially with the development of the internet. The online education paradigm has shifted to collaboration-based knowledge building (Harasim, 2017; Scardamalia & Breiter, 2016). Many studies have proved that well-designed online collaboration learning could improve students' critical thinking, reflection, satisfaction, and attitude (Altınay, 2017; Ku et al., 2013; Lee, 2007; Thompson & Ku, 2006). Collaborative activity promotes initiatives, creativity, critical thinking, co-creation of knowledge, reflection, and transformative learning (Brookfield, 1995; Palloff & Pratt, 2010). Moreover, having a group or community will encourage the exchange of dialogue and reflection (Duffy & Cunningham, 1996).

Asynchronous discussion forums provide students the opportunity to explore their critical thinking in online learning (Beckman & Webber, 2016). The students are encouraged to exercise their higher-order thinking skills into meaningful discussions. Furthermore, they are given time to construct their thoughts and argumentations through reflective thinking in the asynchronous nature of discussion (Garrison et al., 1999). On the other hand, as a facilitator, the teacher plays a pivotal role in directing the discussion through triggers and interventions. Therefore, the transcripts of discussion forums are a high potential to measure the learning engagement in online learning. The use of discussion forums in online learning over the years has resulted in big data in education. These potential data provide education practitioners with essential information in understanding the learning process.

The role of artificial intelligence is very instrumental in exploring big data in education. Recently, research in educational data mining (EDM) and learning analytics (LA) has been rising. Aldowah et al. (2019) classified the applications in these areas into four categories: computer-supported learning analytics, computer-supported predictive analytics, computer-supported behavioral analytics, and computer-supported visualization analytics. Various techniques in EDM and LA, such as prediction, text mining, visualization, association, clustering, and process mining, have been successfully implemented to solve educational problems (Romero & Ventura, 2020). However, Baker (2019) identified the challenges in EDM and LA areas such as transferability, effectiveness, interpretability, applicability, and generalisability. In analysing discussion transcripts, the techniques in text mining plays the leading role. This current study focuses on the recent development of automatic content analysis of asynchronous discussion forum transcripts.

1.1 Content analysis

Krippendorff (2018) defined content analysis as a research method for making reliable, replicable, and valid inferences from texts to the contexts of their use. This scientific method has systematic procedures to investigate the phenomena using specific coding schemes in a text-based environment. The typical phases in the conventional content analysis are theory or rationale, conceptualisation, operationalisation,

coding scheme, sampling, training or pilot reliability, human coding, final reliability, and reporting (Neuendorf, 2002). Furthermore, the types of unit of analysis are message, paragraph, illocution, sentence, and thematic (Rourke et al., 2001; Strijbos et al., 2006). Therefore, content analysis can be used in analysing the various types of educational data. Table 1 below shows the types of data sources used in the content analysis of educational data.

Many previous studies conducted content analysis on various types of data sources such as asynchronous communication (Junus et al., 2019), synchronous communication (Morais & Sampson, 2010), text-based assessment (Chang, 2019; Poldner et al., 2012). Ferreira-Mello et al. (2019) investigated that discussion forum is the most-used data source in analysing text-based educational data between 2016 and 2018. The studies on content analysis of discussion forum transcripts were used to extract meaningful information from such as knowledge construction (Gunawardena et al., 1997), educational experience (Garrison et al., 1999), learning process (Henri, 1992), and critical thinking (Newman, 1995).

The conventional content analysis approach raised issues, such as coping with big data (Garrison et al., 2006). Fahy (2001) also addressed the lack of discriminant capability and reliability. This is because of the complexity of the instrument or the unsuitable unit of analysis. It is also a time-consuming task. Moreover, it cannot provide real-time analysis, which is essential in learning. Manual interpretation is extensively used in the coding process. The coding process requires experienced and trained coders to interpret the data. In addition, learning processes are dynamic, and they change rapidly. Instructors should monitor the discussion routinely in order to maintain the learning engagement. Consequently, there is a need for an automated approach for evaluating texts to understand and improve learning.
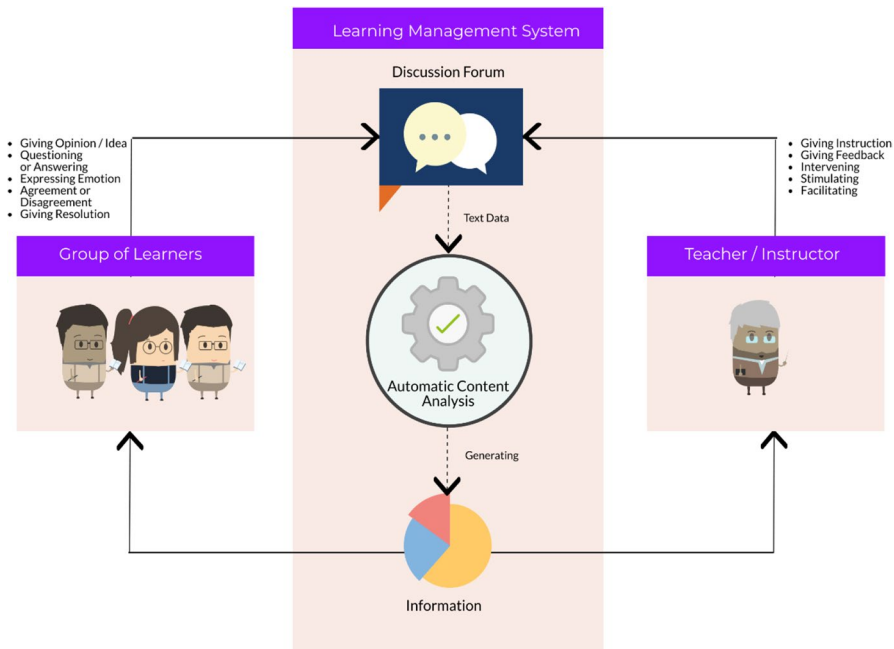
1.2 Automatic content analysis

In recent years, and with the development of technology, the techniques of automatic content analysis have become increasingly advanced. The techniques of text mining play a significant role in analysing educational text automatically. Text mining aims to convert unstructured text into high-quality information (Berry, 2014). Ferreira-Mello et al. (2019) identified four main techniques of text mining in exploring educational data: text classification, natural language processing, information retrieval, text clustering, and summarization.

Figure 1 illustrates the automatic content analysis of online learning discussion forums. Automatic content analysis of discussion transcripts can be used for

Table 1 Type of sources used in content analysis of educational data	Types of data sources	Data sources
	Asynchronous communication	Discussion forum, email
	Synchronous communication	Chat
	Text-based assessment	Online assignment, essay, document, blog

Fig. 1 Overview of automatic content analysis of online discussion forums. Source: Developed by the authors using online design tool Piktochart (https://piktochart.com/)

various aspects of measurement in course resources, learner products of learning, and learner social interactions (Kovanović et al., 2017). On the teacher side, they can also monitor, direct and get the latest information about the dynamics of discussions in forums, such as social presence, cognitive, participation, and engagement. On the learner's side, they can obtain important information in the discussion, such as topics covered, summaries, and level of participation.

The issues that has emerged in the field of content analysis and automatic content analysis are in the aspect of quality (Fahy, 2001; Garrison et al., 2006; Rourke et al., 2001). Neuendorf (2002) defined three standards in evaluating the quality of content analysis procedure: accuracy, precision, and reliability (See Table 2). These quality standards are essential to measure to what extent the performance of automatic content analysis.

Table 2 Quality standards in evaluating content analysis procedures (Neuendorf, 2002)

Standards	Definition
Accuracy	The quality level to which the analysis procedure is bias-free
Precision	The quality level to which the analysis procedure results differences among categories
Reliability	The quality level to which the analysis procedure delivers consistent results on repeated experiments

Recently, there were many systematic literature reviews on educational data mining and learning analytics in general (Aldowah et al., 2019; Koedinger et al., 2015; Romero & Ventura, 2020). In addition, previous studies reviewed the implementation the text mining in education (Ferreira-Mello et al., 2019; Mohammed et al., 2021). However, in recent years, no systematic literature review has explicitly focused on investigating the automatic content analysis of asynchronous discussion forum transcripts. Hence, the current study presents a systematic literature review on the development of automatic content analysis of asynchronous discussion forum transcripts. This article investigates this research area's current state, methods, limitations, and future works between 2016 and 2021.

2 Research questions

The current study aims to explore the current state of the art of researching automatic content analysis in online learning discussion transcripts. Therefore, the following research questions (RQs) guide this study:

RQ1. What kind of measurements were used in analysing online learning discussion transcripts automatically?
RQ2. What theoretical frameworks were used in analysing online learning discussion transcripts automatically?
RQ3. What kind of learning contexts were used to analyse the transcripts of online learning discussions automatically?
RQ4. What kind of methods were used in analysing online learning discussion transcripts automatically?

3 Methodology

According to Jesson et al. (2011), a systematic review must be defined with clearly stated purposes, research questions, and search approaches, stating inclusion and exclusion criteria in producing a qualitative appraisal of the article. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, developed by Page et al. (2021), was our primary guideline in the systematic literature review. PRISMA framework provides comprehensive tools for reporting transparent systematic review such as checklists and flow diagrams. Many studies have used this review guideline in e-learning studies (Jayakumar et al., 2015; Moon & Park, 2021; Peixoto et al., 2021). In addition, the snowballing procedure was added to the PRISMA framework to identify additional relevant studies (Wohlin, 2014). This snowballing procedure aims to complement the results from PRISMA and minimise missing the relevant studies.

3.1 Information sources

This study used five primary electronic academic and citation databases: ACM Digital Library, IEEE Xplore, ScienceDirect, Scopus, and Education Resources Information Centre (ERIC). The publication types of literature reviewed were conferences or journal articles that have gone through a peer-review process. The range of the year of publication of literature is 1 January 2016 to 31 October 2021.

3.2 Eligibility criteria

In conducting the literature review, the eligibility criteria were applied in selecting the relevant articles. These prepared criteria are based on the context and the credibility of the works of literature studied. Table 3 shows the eligibility criteria used as a guide in our study.

3.3 Search strategies

Initially, the relevant studies were searched in five primary electronic databases based on the search string. The search string consisted of three concepts that that related to our research objectives:

1. *The method.* This first concept consisted of several methods of automatic content analysis in analysing text. The keywords are automatic content analysis, text categorisation, text classification, text mining, text segmentation, topic model, and summarisation.
2. *The source of data.* The second concept consisted of the sources of data used in the research. The keyword is asynchronous discussion forum.
3. *The context.* The third concept is the context of the research. Therefore, the keywords are online learning and e-learning.

Based on the three concepts above, our final search string is *('automatic content analysis' OR 'text categorisation' OR 'text classification' OR 'text mining' OR 'text segmentation' OR 'topic analysis' OR 'summarisation) AND ('asynchronous discussion forum') AND ('online learning OR e-learning).*

**Table 3** Eligibility criteria in the current study

Code	Eligibility criteria
IC1	Peer-reviewed
IC2	Conference paper or journal
IC3	Written in English
IC4	The research aims to analyse the discussion transcripts from online discussion forums automatically
IC5	The method is well-defined in the article
IC6	The research context is online learning

3.4 Selection process

In the selection process, there are three main stages: identification, screening, and included. The PRISMA flow chart in Fig. 2 illustrates the article selection process in this study. In identification stage, searches were conducted through five online academic database portals: Scopus, IEEE Xplore, ACM Digital Library, ScienceDirect, and ERIC. The articles that did not meet the eligibility criteria were removed.

The second stage is screening. The first screening is conducted based on the title, abstract, and keywords. Furthermore, we assessed the eligibility of the articles more deeply based on the full text of the article. From this screening stage, 48 articles deserved to be included in this study. The last stage is the finalization of the articles
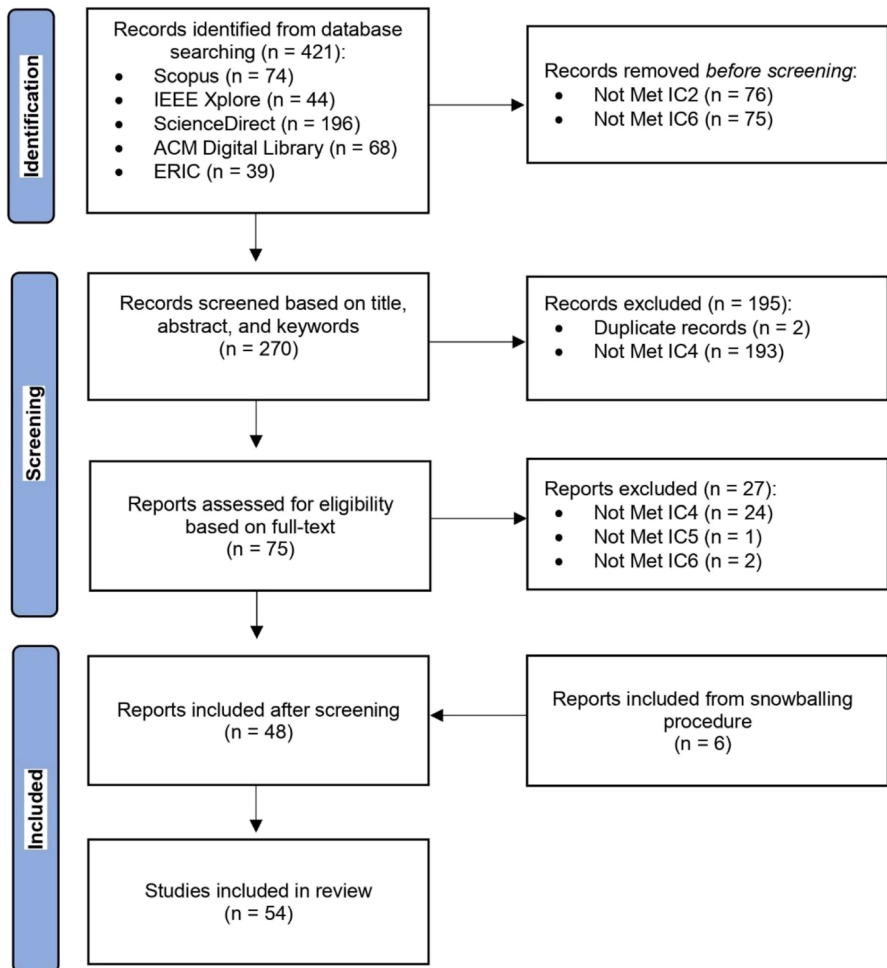
Fig. 2 PRISMA flow chart for selection process. Source: Developed by the authors using PRISMA Framework (Page et al., 2021)

Table 4 Articles included by academic and citation databases

Databases	Publication type		Number of articles
	Conference paper	Journal	
Scopus	4	5	9
IEEE Xplore	11	4	15
ScienceDirect	-	8	8
ACM Digital Library	14	1	15
ERIC	3	4	7
Total	**32**	**22**	**54**
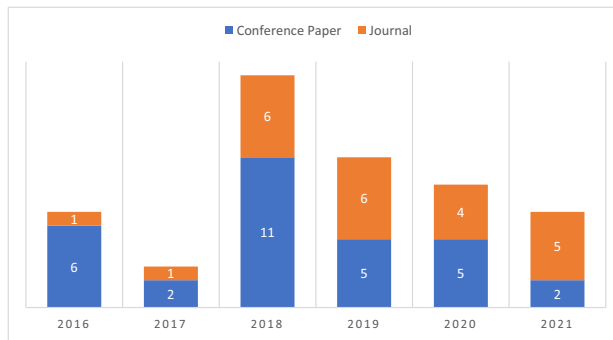
Fig. 3 Articles included by year. Source: Developed by the authors

included in this study. In this case, the snowballing method was used to complement the previous screening process. The snowballing method analyses the articles' references obtained in the previous screening stage to find additional relevant articles. In the end, 54 relevant articles were selected as studies included in our systematic literature review.

4 Findings

The selection process using the PRISMA framework and snowballing procedure resulted in 54 relevant articles. Table 4 shows that most studies were published in conferences proceedings, with 32 articles. Moreover, there are 22 journal articles. IEEE Xplore and ACM Digital Library are the databases that stores the most relevant articles, with 15 articles.

By year of publication, the most articles selected in this literature study were from 2018, with as many as 17 articles (see Fig. 3). There are at least three articles in 2017. It shows that few are still interested in this topic in the early years, namely 2016–2017.

The articles included in the literature study come from credible academic sources in the scientific world that combines technology and education research areas. The

International Conference on Learning Analytics & Knowledge became the publisher that contributed the most to this study, with six articles. This conference is also well known in learning analytics research, and is organised by the Society for Learning Analytics Research (SoLAR). The Internet and Higher Education became the publisher with the most articles, with three articles in journals publications. This journal is ranked first in the field of e-learning, according to SCIMAGO. It shows that this topic is interesting and widely studied by researchers, both in computer and social sciences.

Keyword co-occurrence network was constructed using VOSViewer, as depicted in Fig. 4 (van Eck & Waltman, 2010). Nineteen related keywords were identified with a threshold co-occurrence value is five. These keywords represent the main concepts from included studies. There are three clusters in the co-occurrence network, visualised by different colours.

From Table 5 above, 54 relevant articles were obtained researching automated analysis of learning discussion transcripts in the last six years. Quantitative and qualitative analysis was conducted to dig deeper into several aspects: measurement, data collection, methods, and theoretical frameworks or models used in these
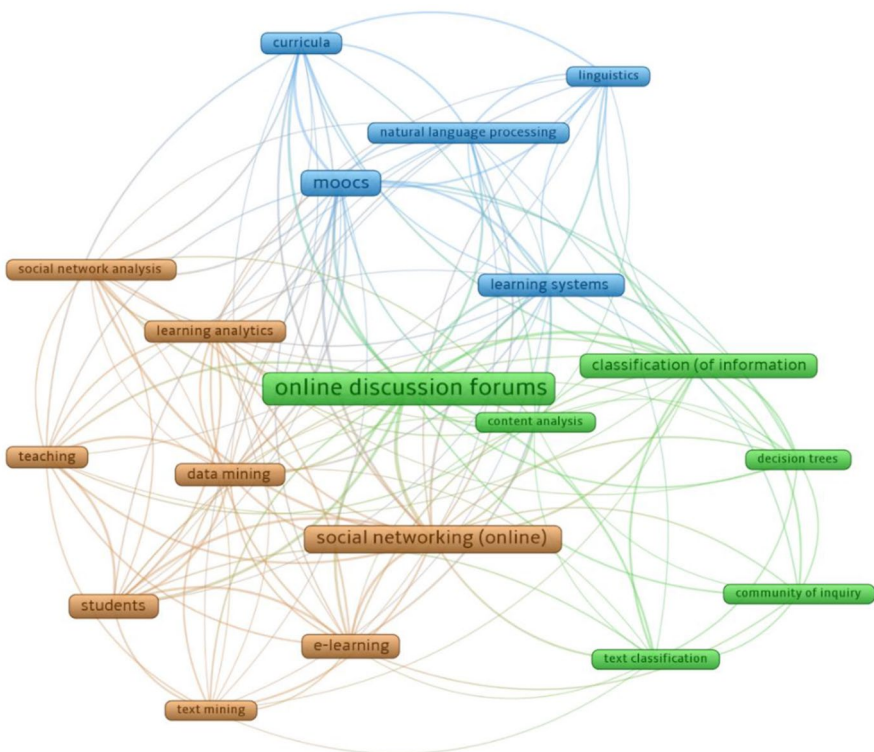
Fig. 4 Visualisation of keyword co-occurrence network. Source: Developed by the authors using VOSViewer (van Eck & Waltman, 2010)

Table 5 Publisher of articles included

Article type	Publisher title	Number of article	Total number of articles
Journal	The Internet and Higher Education	3	22
	Computers & Education	3	
	IEEE Transactions on Learning Technologies	3	
	Computers in Human Behavior	2	
	IEEE Access	1	
	IEEE/ACM Transactions on Networking	1	
	ACM Trans. Comput. Educ	1	
	Education and Information Technologies	1	
	Interactive Learning Environments	1	
	International Journal of Distance Education Technologies	1	
	International Journal of Educational Technology in Higher Education	1	
	Journal of Information Science	1	
	International Journal of Emerging Technologies in Learning	1	
	Frontiers in Psychology	1	
	Data Technologies and Applications	1	
Conference Paper	International Conference on Learning Analytics & Knowledge	7	32
	International Conference of Educational Innovation through Technology (EITT)	2	
	International Conference on Educational Data Mining	2	
	ACM International Conference Proceeding Series	2	
	ACM Conference on Learning at Scale	1	
	ACM Technical Symposium on Computer Science Education	1	
	ACM Symposium on Applied Computing	1	
	Computer Software and Applications Conference (COMPSAC)	1	

Table 5 (continued)

Article type	Publisher title	Number of article	Total number of articles
	Asia–Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data	1	
	Conference of the South African Institute of Computer Scientists and Information Technologists	1	
	European Conference on Technology Enhanced Learning	1	
	Frontiers in Education Conference (FIE)	1	
	IEEE Global Engineering Education Conference (EDUCON)	1	
	IEEE Conference on Computer Communications	1	
	International Conference on Advanced Learning Technologies (ICALT)	1	
	International Conference on Affective Computing and Intelligent Interaction (ACII)	1	
	International Conference on Artificial Intelligence in Education	1	
	International Conference on Information, Intelligence, Systems, and Applications (IISA)	1	
	International Conference on Learning and Optimization Algorithms: Theory and Applications	1	
	International Conference on Learning and Teaching in Computing and Engineering (LaTICE)	1	
	International Conference on Teaching, Assessment, and Learning for Engineering (TALE)	1	
	World Conference on Engineering Education (EDUNINE)	1	
	State-of-the-Art and Future Directions of Smart Learning	1	

studies. Appendix Table 16 shows a summary of selected studies in our literature review.

1. *Measurement.* The measurement can be assessed from the text data of the transcript of the discussion of online learning.
2. *Learning context.* The learning context or scope of previous studies in terms of language and learning mode.
3. *Method.* The methods or techniques used in related studies are based on their approach, algorithm, and results.
4. *The theoretical framework.* The frameworks are used as the theoretical background for these studies.

Figure 5 below illustrates the context study of our systematic literature review in investigating the automatic content analysis of online learning discussion transcripts.

4.1 Addressing RQ1. What kinds of measurements were used in analysing online learning discussion transcripts automatically?

Based on in-depth qualitative analysis, there are eight measurement dimensions in this literature study. The dimensions are cognitive, social, relevance/importance, summary, pattern, behaviour, topic, and pattern. However, these dimensions are not definitive. Hence, there are dependable criteria with these dimensions: objective, variables, key terms, and learning contexts of all selected studies. The summary of the dependable criteria used in analysing the dimensions is in Table 6 below.

Table 7 shows the measurements summary identified from selected studies. The social dimension was the most explored during 2016–2021, with fifteen articles. It shows that social measurement in learning discussions is of great interest to researchers. According to Garrison et al. (1999), social presence is one of the main prerequisites in achieving successful learning discussions. Ferreira et al. (2020) proposed a model that can automatically measure the extent of a learners' social presence (affective, interactive, and cohesive) in a discussion. In addition, Barbosa et al. (2021) explored automatic approach to classify the social presence in Portuguese discussion transcripts.

The social aspect is also closely related to the level of collaboration and interaction of learners in discussions. One of the visualisations to see social interactions in learning discussions is networking. Social network analysis is the most widely used research in this literature study (Boroujeni et al., 2017; Chen et al., 2018; Gašević et al., 2019; Wise & Cui, 2018b; Wise et al., 2017a, 2017b; Xie et al., 2018). In contrast to social networks, epistemic networks visualise the relationships between different concepts used in the discussion text data (Ferreira et al., 2020; Gašević et al., 2019).

The sentiment is also a measurement in the social dimension widely studied by researchers (Fu et al., 2018; Kang et al., 2018; Liu et al., 2018; Moreno-Marcos et al., 2018, 2019). Sentiment analysis is used to measure whether messages or posts
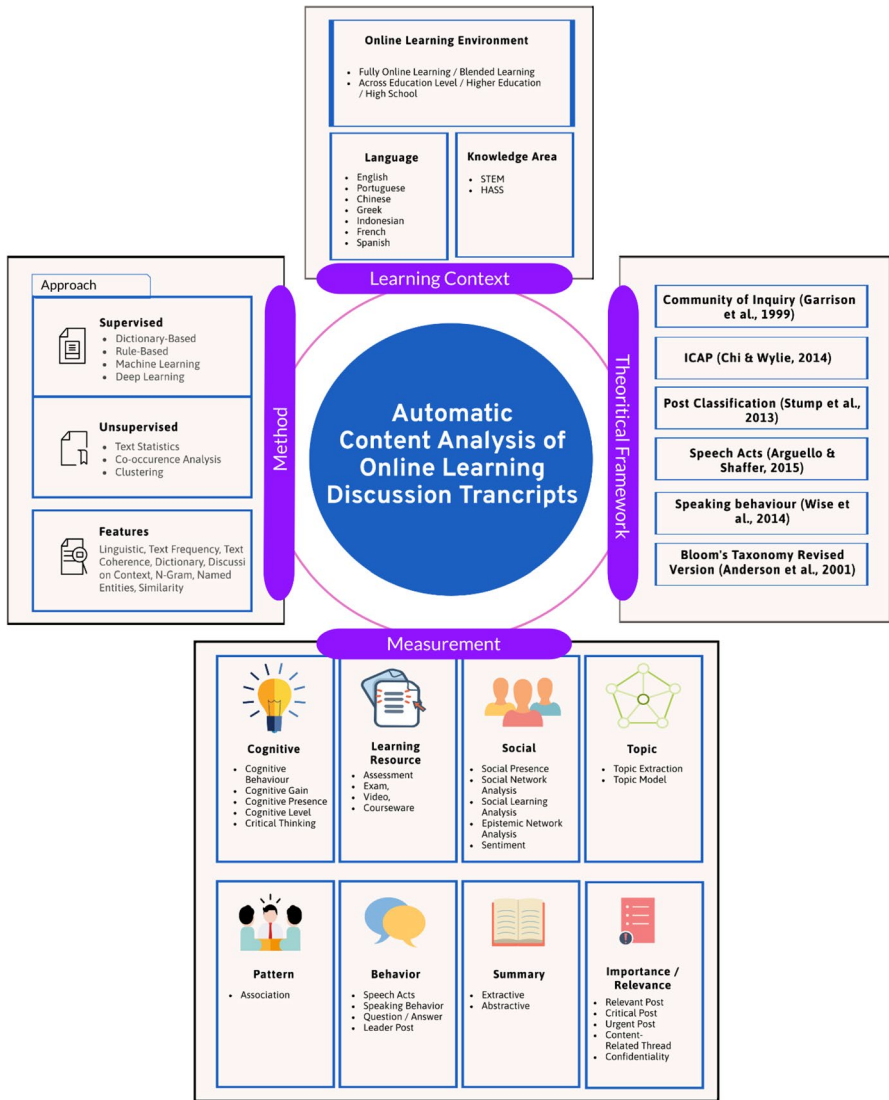
Fig. 5 Automatic content analysis of online learning discussion transcripts review. Source: Developed by the authors using online design tool Piktochart (https://piktochart.com/)

from learners are positive, negative, or neutral. Measurement of this sentiment is important for teachers to see the level of attachment, emotion, and motivation of learners in the ongoing discussion. Thus, teachers can provide fast and appropriate feedback.

The second most studied dimension is relevance or importance. The measurement of the relevance dimension aims to identify the level of relevance or importance of threads, sentences, and posts in the discussion. During much information in

Table 6 Dependable criteria of the measurement dimensions

Dependable criteria	Definition	Examples
Objective	The objective of the study	• Identifying social presence in discussion (Barbosa et al., 2021) • Identifying important post (Machado et al., 2019)
Variables	The variables used in the study	• Affective, interactive, and cohesive (Ferreira et al., 2020) • Important post and not-important post (Sun et al., 2019)
Key Terms	The key terms or keywords found in the study	• Community of inquiry, content analysis, content analytics, online discussion, text classification (Kovanovic et al., 2017) • Computer-mediated communication, navigation, MOOC (Almatrafi et al., 2018)
Learning Contexts	The learning contexts of the study	• Higher education (Neto et al., 2018) • MOOC (Moreno-Marcos et al., 2019)

Table 7 Measurement of selected studies in literature review

Dimensions	Measurements	Variables	Studies	Total
Social	Social presence	• Affective • Interactive • Cohesive	(Barbosa et al., 2021) (Ferreira et al., 2020) (Zou et al., 2021)	15
	Social network	• Density • Centrality • In-degree • Out-degree • Sociogram	(Boroujeni et al., 2017) (Chen et al., 2018) (Gašević et al., 2019) (Kang et al., 2018) (Wise & Cui, 2018b) (Wise et al., 2017a, 2017b) (Xie et al., 2018)	
	Social learning network	• Neighbourhood-based • Path-based • Post-based	(Yang et al., 2018)	
	Epistemic network	• Projection graph • Network graph • Subtraction network	(Gašević et al., 2019) (Ferreira et al., 2020)	
	Sentiment	• Positive • Negative • Neutral	(Fu et al., 2018) (Kang et al., 2018) (Liu et al., 2018) (Moreno-Marcos et al., 2018) (Moreno-Marcos et al., 2019)	
Relevance or Importance	Relevance of thread	• Content-related thread • Non-content-related thread	(Wise et al., 2016) (Wise & Cui, 2018a) (Wise & Cui, 2018b) (Wise et al., 2017a, 2017b) (Wise et al., 2017a, 2017b)	13

Table 7 (continued)

Dimensions	Measurements	Variables	Studies	Total
	Importance of post	• Important post • Not important post	(Almatrafi et al., 2018) (Liu et al., 2016) (Machado et al., 2019) (Sun et al., 2019)	
	Importance of sentence	• Important sentence • Unimportant sentence	(Le et al., 2018)	
	Type of thread	• Meetups • General discussions • Lectures • Assignments • Platform • Issues and course feedback	(Feng et al., 2018)	
	Relevance of starting post	• Content-related • Logistics-related • No-action-required	(Ntourmas et al., 2019) (Ntourmas et al., 2021)	
Cognitive	Cognitive presence	• Triggering event • Exploration • Integration • Resolution • Other	(Barbosa et al., 2020) (Barbosa et al., 2021) (Farrow et al., 2019) (Hayati et al., 2019) (Kovanovic et al., 2017) (Neto et al., 2018) (Yang et al., 2022)	12
	Cognitive gain	• Cognitive gain	(Wu & Wu, 2018) (Huang et al., 2019)	
	Metacognition	• Confident • Unconfident • Either		
	Cognitive processing	• Analytical thinking • Clout • Authenticity	(Moore et al., 2019)	

Table 7 (continued)

Dimensions	Measurements	Variables	Studies	Total
	Cognitive engagement	•Active •Constructive •Interactive •Passive	(Wu & Wu, 2018) (Hayati et al., 2020)	
	Cognitive level	•Remembering •Understanding •Applying •Analysing •Evaluating •Creating	(Cheng et al., 2021)	
Topic	Topic model	•Topic extraction •Topic ranking •Visualisation	(Atapattu et al., 2016) (Brinton et al., 2018) (Chen et al., 2021) (Gottipati et al., 2019) (Liu et al., 2018) (Peng et al., 2020) (Peng et al., 2021) (Rolim et al., 2019) (Setiawan et al., 2020) (Wong et al., 2016) (Yang et al., 2018) (Zarra et al., 2018)	12
Behaviour	Speech acts of post	•Information seeking •Information giving •Other	(Boroujeni et al., 2017) (Hecking et al., 2016) (Joksimovic et al., 2020)	9
	Confidential information	•Confidential entity •Non-confidential entity	(Bosch et al., 2020)	

Table 7 (continued)

Dimensions	Measurements	Variables	Studies	Total
	Speaking behaviour of post	• Argumentation • Responsiveness • Elicitation • Reflection on individual process • Reflection on group process	(Chen et al., 2018)	
	Question / answer	• Question • Answer	(Brinton et al., 2018) (Irish et al., 2020) (Rolim et al., 2019)	
	Leader post	• Leader post • Non-leader post	(Xie et al., 2018)	1
Learning Sources	Type of learning resource	• Assessment • Exam • Video • Courseware ('Read, 'Slides', 'Transcript, and 'Additional resources')	(An et al., 2019)	1
Summary	Summary	• Topic-based summary	(Gottipati et al., 2019)	1
Pattern	Chance discovery	• Keygraph	(Wong & Li, 2016)	1

discussion forums, this automated analysis is helpful for students in finding which threads are related to the substance of learning (Wise & Cui, 2018a, 2018b; Wise et al., 2016, 2017b, 2017a). This level of importance measurement can be divided into two forms, namely sentences (Le et al., 2018) or posts (Almatrafi et al., 2018; Liu et al., 2016; Machado et al., 2019; Sun et al., 2019).

The next most explored dimension is cognitive. The closest dimension to achieving the main goal in the learning discussion, namely acquiring new knowledge. According to the Community of Inquiry (CoI) framework (Garrison et al., 1999), the degree to which learners construct meaning through reflection and critical conversation is represented by cognitive presence. Therefore, automatic cognitive presence analysis can help teachers measure the quality of discussion quickly (Barbosa et al., 2020; Barbosa et al., 2021; Farrow et al., 2019; Hayati et al., 2020; Hayati et al., 2019; Kovanovic et al., 2016; Neto et al., 2018). Furthermore, Yang et al. (2022) used this cognitive presence to measure the critical thinking level of the students: low, medium, and high.

The measurement of cognitive engagement has also been extensively researched to see how learners feel attached to the activities in the discussion (Hayati et al., 2020; Wu & Wu, 2018). This cognitive engagement was measured using the ICAP framework with interactive, constructive, active, and passive modes of behaviour (Chi & Wylie, 2014). Cognitive processing was also investigated based on analytical thinking, influence/clout, and authenticity (Moore et al., 2019). In addition, Huang et al. (2019) proposed an automatic analysis method to distinguish students' confident or unconfident sentences.

Topic measurement is also one of the trends in discussion transcript analysis research (Atapattu et al., 2016; Brinton et al., 2018; Liu et al., 2018; Peng et al., 2020, 2021; Rolim et al., 2019; Wong et al., 2016; Yang et al., 2018; Zarra et al., 2018). Extracting and modelling this topic helps see what topics are being discussed in the forum. In addition, this measurement is also often combined with other measurements such as summarisation. Gottipati et al. (2019) proposed a method of summarising the content of discussion based on a topic model. However, unfortunately, there is very little research on auto-summarising the content of discussions in recent years.

Some of the measurements analysed in the dimensions of learner behaviour are speech acts, speaking behaviour, leadership, and questions or answers. Speech act analysis aims to categorise discussion messages into seeking information or providing information (Boroujeni et al., 2017; Hecking et al., 2016; Joksimovic et al., 2020). This analysis is similar to identifying the form of a message in the form of a question or answer (Brinton et al., 2018; Irish et al., 2020; Rolim et al., 2019). Chen et al. (2018) also analysed the lexical features in identifying speaking behaviour in a message. Xie et al. (2018) identified learners who become leaders based on the messages conveyed in the discussion. Bosch et al. (2020) also proposed a solution to hiding confidential information in discussions with an automated classification process.

Pattern measurement is also an interesting aspect to study. This measurement is closely related to prediction. Wong and Li (2016) analysed patterns in discussions with chance discovery analysis to predict students' academic performance. However,

the research in pattern prediction with the concept of serendipity learning is still limited. In addition, An et al. (2019) proposed a method to link text in discussion with resources in online learning.

These eight dimensions measurements above show the research trend in automatic content analysis of asynchronous discussion forum transcripts. These dimensions indicate different research focus. Fifty-nine percent of studies focus on helping teachers understand learning processes in the discussion forum. These learning processes are measured from social, cognitive, and behaviour dimensions. These three dimensions are vital in constructing knowledge and improving critical thinking through meaningful discussions. Therefore, the teachers should obtain the information to provide appropriate feedback. These measurements are generally difficult to perform compared to others due to their implicit nature. Table 8 below shows the trend of studies.

Ten studies aim to measure the primary information from discussion to help students learn better from discussions. The automatic topic extraction and summary measurements are helpful for students to understand the discussions quickly. Most studies also developed supporting tools to detect and visualise the topic and summary. The other studies focus on helping the participants conduct the discussions better with relevance and learning resources measurements. These measurements help the participants organize and select the related items in discussions such as posts, threads, and sentences. In addition, one study focuses on predicting students' performance in pattern measurement.

4.2 Addressing RQ2. What theoretical frameworks were used in analysing online learning discussion transcripts automatically?

The theoretical framework is fundamental in content analysis research. Analysis of online learning discussion transcripts certainly requires guidance, especially in classifying or categorising messages, threads, or sentences. In the current study, there are some criteria used in including the theoretical frameworks. The theoretical framework must be equipped with a coding scheme or basic guideline for analysing text. It also reliable and valid framework. It can be showed from the impact of the theoretical framework.

Table 8 The trend of studies based on the measurement dimensions

Dimensions	Focus of Studies	User	Total Studies
•Social •Cognitive •Behaviour	Understanding learning presence of students in the discussion	Teacher	32
•Relevance or importance •Learning resources	Helping participants in organizing and navigating the discussion	Teacher Student	14
•Topic •Summary	Extracting main points from discussion transcripts	Teacher Student	10
•Pattern	Predicting students' performances in the discussion	Teacher	1

Therefore, the number of citations of original paper should be high. There are six theoretical framework identified: Community of Inquiry (Garrison et al., 1999), Post Classification (Stump et al., 2013), Speech Acts (Arguello & Shaffer, 2015), ICAP (Chi & Wylie, 2014), Speaking behaviour (Wise et al., 2014) and Bloom's Taxonomy Revised Version (Anderson et al., 2001). However, there is only 41% studies used these theoretical frameworks. Table 9 shows which frameworks have been used in recent years.

As Table 9 shows, the CoI framework is the most-used framework for content analysis for a text-based learning environment. This framework consists of three prerequisite elements for achieving successful discussion: cognitive, social, and teaching presences. Six studies explored automated cognitive presence measurement approaches (Barbosa et al., 2020; Farrow et al., 2019; Hayati et al., 2019; Hayati et al., 2020; Kovanovic et al., 2017; Neto et al., 2018). Ferreira et al. (2020) also proposed a method for analysing social presence. However, over the last six years, no one has researched an automated approach to measuring teaching presence.

Another popular theoretical framework used in analysing cognitive engagement is the ICAP framework (Chi & Wylie, 2014). The ICAP framework has four modes, namely Interactive, Constructive, Active, and Passive. These four modes show the learner's behaviour or cognitive engagement in a learning discussion. Wu and Wu (2018) analysed cognitive behaviour and used it to measure predicted learning gain. In addition, Hayati et al. (2020) also investigated learners' cognitive engagement and social interaction in discussion.

On the other hand, Stump's framework has been used by five studies (Stump et al., 2013). This framework has three main categories of messages, namely seeking information, providing information, and others. It was used by the same first author (Wise & Cui, 2018a, 2018b; Wise et al., 2016, 2017b, 2017a). The Speech Acts framework is also used to identify questions, answers, issues, issue resolutions, positive acknowledgements, and negative acknowledgements in a message (Arguello & Shaffer, 2015). Hecking et al. (2016) used it to see the level of information exchange in discussions. Boroujeni et al. (2017) used this framework to predict discussion activity from the content dimension. In addition, this framework is also used to analyse the dynamics of the implementation of the discussion (Joksimovic et al., 2020).

Speaking behaviour in discussion is also used as the basis for research related to analysis in the discussion. Wise et al. (2014) proposed a framework in collaborative learning on how to explore listening (message access) and speaking (message writing) activities in online discussions. Chen et al. (2018) combined this framework with social network analysis to generate insights for discussion participants. In addition, Bloom's taxonomy also used as measurements of cognitive skills in discussion. Cheng et al. (2021) proposed automatic evaluation of cognitive level based on deep learning network.

4.3 Addressing RQ3. What kind of learning contexts were used to analyse the transcripts of online learning discussions automatically?

The context of learning is very influential in research in the field of education. The results obtained in research related to online learning discussions cannot be

Table 9 Theoretical frameworks of selected studies in literature review

Theoretical framework	Elements	Description	Studies	Total studies	Total citation*
Community of Inquiry (Garrison et al., 1999)	Cognitive presence	The presence indicates the extent to which learners construct meaning through reflection and critical conversation in discussions	(Barbosa et al., 2020)(Barbosa et al., 2021)(Farrow et al., 2019)(Hayati et al., 2019)(Kovanovic, et al., 2017)(Neto et al., 2018)(Neto et al., 2021)(Yang et al., 2022)	10	7,392
	Social presence	The presence shows the ability of participants to present themselves socially and emotionally in discussions	(Barbosa et al., 2021)(Ferreira et al., 2020)(Zou et al., 2021)		
	Teaching presence	The presence shows the teacher's role in directing and guiding the discussion	-		
Post Classification (Stump et al., 2013)	Information seeking	The message from learners aims to seek help, information, or direction	(Wise & Cui, 2018a)(Wise & Cui, 2018b)(Wise et al., 2016)(Wise et al., 2017a, 2017b)(Wise et al., 2017a, 2017b)	5	63
	Information giving	The message from learners aims to provide help or information			
	Other	The message that does not explicitly seek or provide information			
Speech Acts (Arguello & Shaffer, 2015)	Question	The message asking for information in the discussion	(Boroujeni et al., 2017)(Hecking et al., 2016)(Joksimovic et al., 2020)	3	47

Table 9 (continued)

Theoretical framework	Elements	Description	Studies	Total studies	Total citation*
	Answer	The message that answers questions and provides information in discussions			
	Issue	The message that conveys dissatisfaction or difficulties in learning			
	Issue resolution	The message that responds to and resolves the issues raised			
	Positive acknowledgement	The message that expresses a positive sentiment			
	Negative acknowledgement	The message that expresses a negative sentiment			
ICAP (Chi & Wylie, 2014)	Interactive	Interact and collaborate with colleagues	(Wu & Wu, 2018) (Hayati et al., 2020)	2	1,338
	Constructive	Productive in providing ideas and building knowledge			
	Active	Participate, understand the discussion, and focus on learning activities			
	Passive	Receive information only in discussions without doing other activities related to the learning process			
Speaking behaviour (Wise et al., 2014)	Argumentation	The message conveyed by the learner is in the form of an argument/idea accompanied by evidence	(Chen et al., 2018)	1	89

Table 9 (continued)

Theoretical framework	Elements	Description	Studies	Total studies	Total citation*
	Responsiveness	Messages conveyed to respond to ideas			
	Elicitation	The direction of the questions in the discussion			
	Reflection on individual process	Messages that contain reflections in the individual learning process			
	Reflection on group process	Messages that contain reflections in the group learning process			
Bloom's taxonomy revised version (Anderson et al., 2001)	Remembering	The message that conveys the basic information from students' long-term memory	(Cheng et al., 2021)	1	25,461
	Understanding	The message that conveys the elaboration process to extract the meaning of the information			
	Applying	The message that conveys the implementation of specific ideas			
	Analysing	The message that conveys the connections process among ideas			
	Evaluating	The message that conveys the judgment of the ideas			
	Creating	The message that conveys the formulation process of the new solution or model			

*= total citations from Google Scholar (10 November 2021)

separated from the type of learning, the nature and the language used in the dataset. Each language has a different vocabulary, rules, and lexical. In recent years, seven languages have been used in the selected studies (see Table 10).

English has become the most studied language in the last six years, with 42 studies. The recent development indicates that text mining of English text is more advanced than other languages. Many datasets, dictionaries, tools, and libraries in English help researchers study discussion transcripts such as LIWC, Stanford CoreNLP, Coh-Metrix, and SpaCy. The general approach in analysing text is to use features or dictionaries found in the same language. However, there is a specific approach utilising other languages to analyse discussion transcripts. This approach uses datasets with multiple languages in the experiment. Table 11 below shows the studies that used the cross-language approach.

Three studies proposed the cross-language-based approach. Feng et al. (2018) used several language features to categorise threads in English, French, Chinese, and Spanish. Barbosa et al. (2020) used a model trained by English data to classify transcripts of Portuguese discussions. In addition, Barbosa et al. (2021) proposed an automatic text translation method on English and Portuguese datasets to identify social and cognitive presences. The experiment improved performance when using English translation (Portuguese translated to English). In contrast to Portuguese translation, the performance was decreased. It shows that English has more significant linguistic features than other languages in classifying messages of discussion transcripts.

In terms of the types of online learning, we found two types examined in the research in this literature study: fully online and blended learning. However, one more category was added, because fourteen articles do not clearly describe their type of online learning. Fully online learning is the most widely studied. It could be because this type of learning will result in extensive discussion data. Generally, more data will improve the performance of the algorithm. Fully online learning has two educational level contexts: across education level and higher education. There is also one study in high school. On the other hand, there has been little research about automatic content analysis from blended learning. Blended learning mainly uses online discussion as a supporting tool. It is quite different from fully online learning, where discussion forums are one of the primary communication tools. Table 12 shows the type of online learning of selected studies in this literature review.

Table 10 Languages of dataset in selected studies	Languages	Total
	English	42
	Chinese	7
	Portuguese	5
	Greek	2
	Indonesian	1
	French	1
	Spanish	1

Table 11 Cross-language studies

Language of training dataset	Language of testing dataset	Method	Study
English	•French •Spanish •Chinese	Gradient Boosting Decision Tree (GBDT) with Linear Regression (LR)	(Feng et al., 2018)
English	•Portuguese	Random Forest	(Barbosa et al., 2020)
•English (Original) •English (Translation from Portuguese) •Portuguese (Original) •Portuguese (Translation from English)	•English (Original) •English (Translation from Portuguese) •Portuguese (Original) •Portuguese (Translation from English)	Random Forest and Automatic Text Translation	(Barbosa et al., 2021)

Based on knowledge area perspective, there are two categories of datasets: science, technology, engineering, and mathematics (STEM), and humanities, arts, and social sciences (HASS). These terms are common knowledge areas broadly used in the educational curriculum of many countries (Gonzalez & Kuenzi, 2012; Turner & Brass, 2014). These knowledge areas should be acknowledged when conducting automatic content analysis as the nature of the research. Table 13 shows the comparison of the knowledge areas used in datasets.

From the data in Table 13, it is apparent that 59% of studies used single-disciplinary datasets. Single-disciplinary researches only used either STEM or HASS courses as datasets. STEM courses have been used more than HASS's with 27 studies. The STEM datasets used are programming, machine learning, software engineering, statistics, medicine, information systems, artificial intelligence, logic, technology, and discrete mathematics. On the other hand, the HASS's datasets are psychology, history, law, general education, and English literature. These single-disciplinary researches also pointed out their limitations in generalisability issues.

Table 12 Type of online learning of selected studies in literature review

Types of online learning	Context	Total
Fully online learning	Across education levels or Massive Open Online Courses (MOOCs)	26
	Higher education	13
Blended learning	Higher education	1
Not reported	High school	1
	Higher education	13

Table 13 Knowledge areas of courses used as datasets

Categories	Courses' knowledge areas	Total
Single-disciplinary	STEM	27
	HASS	5
Multidisciplinary	STEM and HASS	14

In multidisciplinary researches, fifteen studies used both STEM and HASS courses as datasets in their experiments. These studies focussed on generalisability issues in analysing discussion transcripts with different knowledge areas. Wise et al., (2017a, 2017b) proposed a model to detect content-related threads on the cross-domain datasets from statistics and psychology courses. However, the generalisability experiments show poor results. Similar unsatisfactory results were also obtained from other studies. Almatrafi et al. (2018) used medicine and education courses as datasets on identifying urgent posts. Ntourmas et al. (2019) also reported non-acceptable results when testing their model on python programming and history courses.

These unsatisfactory results could be due to the differences in distribution and the unbalanced nature of datasets (Neto et al., 2021). Farrow et al. (2019) emphasised rebalancing classes and avoiding data contamination steps to improve generalisability. These findings indicate that generalisability is the main challenges of automatic content analysis that needs further investigation, especially cross-domain context. In addition, there were significant differences between STEM and HASS courses (Bates, 2015). Hence, the researchers need to define this context in their content analysis studies.

4.4 Addressing RQ4. What kind of methods were used in analysing online learning discussion transcripts automatically?

In recent years, previous studies used machine learning, rule-based, text-statistics, and co-occurrence methods in analysing online learning discussion transcripts automatically. Ninety-three percent of studies used the machine learning method. It can be concluded that machine learning still plays the leading role in this research area. Table 14 below shows a summary of the methods used in the selected studies.

There are two approaches in automated analysis of text data in online learning discussions, namely supervised and unsupervised. The supervised approach uses pre-labelled datasets. Therefore, training data needs to be prepared by giving labels or classes in advance. This data labelling process can be carried out based on the assessment of several coders manually or automatically. The supervised machine learning method performs training on sets where the algorithm learns based on the provided data set to get predictive features. Next, the model is evaluated by validating and testing with different data sets. Thirty-five studies explored the supervised machine learning methods. The common problems in the supervised approach are classification and regression modelling. In addition, five studies used deep learning methods. Deep learning imitates the concept of the human brain working with

Table 14 The methods of selected studies in literature review

Methods	Approaches	Algorithms	Studies	Total
Machine Learning	Supervised	Logistic Regression	(Almatrafi et al., 2018) (Wise & Cui, 2018a, 2018b) (Wise et al., 2016) (Wise et al., 2017a, 2017b) (Wise et al., 2017a, 2017b) (Wu & Wu, 2018) (Xie et al., 2018)	50
		Linear Regression (LR)	(Feng et al., 2018)	
		Random Forest	(Almatrafi et al., 2018) (Barbosa et al., 2020) (Barbosa et al., 2021) (Boroujeni et al., 2017) (Farrow et al., 2019) (Ferreira et al., 2020) (Hecking et al., 2016) (Kovanovic, et al., 2017) (Liu et al., 2016) (Moreno-Marcos et al., 2018) (Moreno-Marcos et al., 2019) (Neto et al., 2018) (Neto et al., 2021)	
		Naïve Bayes	(Atapattu et al., 2016) (Hayati et al., 2019) (Le et al., 2018)	
		Gradient Boosting Decision Tree (GBDT)	(Feng et al., 2018)	
		Extra-Trees	(Bosch et al., 2020)	
		Weighted Decision Tree	(Ntourmas et al., 2021)	

Table 14 (continued)

Methods	Approaches	Algorithms	Studies	Total
		Adaptive Boosting Model (AdaBoost)	(Almatrafi et al., 2018) (Xie et al., 2018)	
		Hidden Markov Models (HMM)	(Joksimovic et al., 2020)	
		Support Vector Machines (SVM)	(Hayati et al., 2020) (Ntourmas et al., 2019) (Rolim et al., 2019)	
		Long Short-Term Memory–Convolutional Neural Network (LSTM-CNN)	(Yang et al., 2022)	
		Long Short-Term Memory–Recurrent Convolutional Neural Networks (LSTM-RCNN)	(Sun et al., 2019)	
		Long Short-Term Memory–Conditional Random Field (LSTM-CRF)	(An et al., 2019)	
		Deep Neural Network (DNN)	(Bosch et al., 2020)	
		Attention Mechanism, Bidirectional Gated Recurrent Unit, and Convolutional Neural Network Model (AM-BiGRU-CNN)	(Cheng et al., 2021)	
		Bidirectional Encoder Representations from Transformers (BERT)	(Zou et al., 2021)	
	Unsupervised	Latent Dirichlet Allocation (LDA)	(Atapattu et al., 2016) (Brinton et al., 2018) (Chen et al., 2021) (Gašević et al., 2019) (Joksimovic et al., 2020) (Peng et al., 2020) (Peng et al., 2021) (Rolim et al., 2019) (Wong & Li, 2016) (Wong et al., 2016) (Yang et al., 2018) (Zarra et al., 2018)	
		Latent Semantic Analysis (LSA)	(Gottipati et al., 2019) (Setiawan et al., 2020)	
		K-means	(Gottipati et al., 2019)	

Table 14 (continued)

Methods	Approaches	Algorithms	Studies	Total
Rule-Based	Supervised	Dictionary-Based	(Barbosa et al., 2020) (Barbosa et al., 2021) (Chen et al., 2018) (Farrow et al., 2019) (Ferreira et al., 2020) (Huang et al., 2019) (Kang et al., 2018) (Kovanovic, et al., 2017) (Liu et al., 2018) (Machado et al., 2019) (Moore et al., 2019) (Moreno-Marcos et al., 2018) (Moreno-Marcos et al., 2019) (Neto et al., 2018) (Peng et al., 2020) (Peng et al., 2021)	16
Text Statistics	Unsupervised	Tf-idf	(Irish et al., 2020)	1
Co-occurrence Analysis	Unsupervised	Dense subgraphs	(Fu et al., 2018)	1

artificial neural networks. Deep learning is generally used to solve complex problems with extensive data. However, these studies are relatively new, between 2019 and 2021.

Unsupervised machine learning used to solve clustering problems. This approach analyses text that has no class or label. The algorithm measures the level of similarity or difference to group data, such as the extraction of topics discussed in a discussion. Topic modelling is a research problem that has been widely studied in recent years with fourteen studies. This automatic topic extraction is beneficial for both students and teachers to see the main conversation in the discussion. The most widely used algorithm for modelling topics in this literature study is Latent Dirichlet Allocation (LDA) with twelve studies.

In general, topic modelling methods were used to obtain the main topics of discussions. Recently, these methods were implemented to develop tools in detecting and visualising topics (Atapattu et al., 2016; Wong et al., 2016; Zarra et al., 2018). Also, these methods were combined with other methods to achieve different objectives. Brinton et al. (2018) investigated whether a student asks or answers a question using the topic modelling aproach. Rolim et al. (2019) identified students' weaknesses and strengths using LDA and SVM. Gottipati et al. (2019) developed a tool to generate a summary based on the topics from online learning discussion transcripts. Gašević et al. (2019) presented a method to analyze collaborative learning using a topic model, social network analysis, and epistemic network analysis. Three studies proposed models to understand forums with time, emotions, and topics features (Liu et al., 2018; Peng et al., 2020, 2021). We predict that the methods-combining approach will be essential in the future. The topic modelling methods can be used in other perspectives. The trend shows that LDA is a fundamental method in topic modelling. This method has the potential to develop further. Three studies developed new methods based on LDA (Liu et al., 2018; Peng et al., 2020, 2021).

Random forest is also predicted to be essential in the future. This algorithm is the most popular in solving the text classification problem with thirteen studies. Random forest is based on random decision trees. This ensemble algorithm is relatively faster to train than the others (Kowsari et al., 2019). Moreover, it also reduces variance. The performance of this algorithm is proven to be good in analysing discussion transcripts (Barbosa et al., 2021; Farrow et al., 2019; Ferreira et al., 2020; Kovanovic et al., 2017; Neto et al., 2021). However, special attention is needed in selecting decision trees to avoid overfitting. Too many decision trees used will increase complexity.

Dictionary or rule-based methods have different concepts from machine learning or deep learning. This method uses a guide that has been formulated previously, namely a dictionary or rules. Sixteen studies used this approach using existing dictionaries such as LIWC, Coh-Metrix, and NLPIR. The dictionary-based method is becoming an increasingly attractive method for the future because it provides advantages in terms of process speed and collaboration in dictionary development (Scharkow, 2017). Combining machine learning with dictionary or rule-based methods is also a promising approach in examining texts in discussion transcripts.

Other methods used are text statistics and co-occurrence analysis. These are the earliest methods used in the introduction of automatic content analysis. The text

statistics uses statistical measurement in analysing discussion transcripts automatically. Irish et al. (2020) used the term frequency and the inverse document frequency (tf-idf) to find similar questions in discussion forums. In addition, co-occurrence analysis is an extension of text statistics method. This analysis uses the co-occurrence of specific features in the discussion transcripts such as word frequency, semantic, and time to construct association. Fu et al. (2018) used the co-occurrence analysis to group users in discussion forums with dense subgraphs visualization.

From the discussion above, each method provides different approaches and characteristics. We suggest combining machine learning and rule-based methods. Eleven studies proved that this combining method resulted in good performances (Barbosa et al., 2020; Barbosa et al., 2021; Farrow et al., 2019; Ferreira et al., 2020; Kovanovic et al., 2017; Liu et al., 2018; Moreno-Marcos et al., 2018; Moreno-Marcos et al., 2019; Neto et al., 2018; Peng et al., 2020, 2021). The experiments show that dictionaries features are also good predictors. Moreover, the researchers should implement more than one classifier in their experiments. This suggestion aims to obtain the best model based on the evaluation results.

The latest approaches in natural language processing will be more significant in the future, such as BERT (Devlin et al., 2019) and GPT-3 (Brown et al., 2020). These fine-tuning approaches use pre-trained language models with hyper-parameters in understanding text. Zou et al. (2021) showed that BERT performed promising results with accuracy, precision, and recall above 0.80. However, no study used GPT-3 to explore discussion transcripts automatically. Studying how far these latest algorithms will develop in this research area will be fascinating.

All of these studies are experimental. In machine learning, the standard evaluation measurements are kappa statistic, accuracy, precision, recall, and F-measure. These also represent the quality of automatic content analysis. The experiments' results in Table 15 shows that the performance of the previous studies sounds promising. Nineteen studies reported high performance with an accuracy level above 80%. It also shows good performance in other evaluation metrics: precision, recall, and F-measure. However, there are no studies that reported kappa statistics more than 80%. Seventeen studies reported kappa statistics results that range from moderate to substantial levels of agreement. The kappa statistics are used to measure the reliability of the content analysis (Landis & Koch, 1977). These results indicate that there is an excellent opportunity for future research to improve reliability.

It should be noted that most studies used different datasets and learning contexts. Therefore, it cannot be compared directly. These results should be related to the contexts of the datasets. The detail of the experiments' results can be seen in Appendix Table 16.

5 Discussions, Limitations and Research Directions

Insightful findings were found in the systematic literature review for future research in analysing discussion transcripts automatically. There are eight dimensions of measurement found in previous studies. These dimensions are cognitive, social, relevance/interest, summary, pattern, behaviour, topic, and learning resources. The top

Table 15 The experiments' results in selected studies

The evaluation metrics	The range of value	Total studies
Kappa Statistic	≤ 0.40	-
	0.41 – 0.60	8
	0.61 – 0.80	9
	0.81 – 1.00	-
Accuracy	≤ 0.40	-
	0.41 – 0.60	-
	0.61 – 0.80	9
	0.81 – 1.00	19
Precision	≤ 0.40	-
	0.41 – 0.60	-
	0.61 – 0.80	6
	0.81 – 1.00	9
Recall	≤ 0.40	-
	0.41 – 0.60	2
	0.61 – 0.80	10
	0.81 – 1.00	5
F-measure (F1)	≤ 0.40	-
	0.41 – 0.60	3
	0.61 – 0.80	10
	0.81 – 1.00	8

three dimensions that are widely studied are social, relevance/importance, and cognitive. The theoretical framework guides researchers in analysing these dimensions. Six theoretical frameworks were identified, namely Community of Inquiry (Garrison et al., 1999), Post Classification (Stump et al., 2013), Speech Acts (Arguello & Shaffer, 2015), ICAP (Chi & Wylie, 2014), Speaking behaviour (Wise et al., 2014), and Bloom's Taxonomy Revised Version (Anderson et al., 2001). The CoI is the framework studied the most, in as many as ten studies. It suggests that the CoI framework can serve as an excellent comprehensive guide in measuring educational experiences in text-based environment.

The context of learning is vital in research in the world of education. This study found that most studies examine fully online learning, whether across education levels (MOOCs) or higher education levels. Fully online learning that can generate much data is one of the advantages for researchers. However, blended learning should also receive special attention for future research. Also, selected studies used seven languages: English, Chinese, Portuguese, Indonesian, French, Spanish, and Greek. The majority of research explores English, which is 78%. It opens opportunities for researchers to delve deeper into other languages. Furthermore, advanced research on English allows it to be combined with other languages with a cross-language approach.

There are two approaches used in research related to automatic learning discussion transcripts, namely supervised and unsupervised. This approach aims to solve classification (thread, message, and sentence) and prediction problems in discussion. On the other hand, an unsupervised approach is used for modelling topic and message associations in online learning discussions. In our selected studies, machine learning is the favoured method to evaluate texts in discussion forums, in both

supervised and unsupervised tasks. Out of 54 studies, 38 used machine learning as their primary method. In addition, nine studies combine machine learning with dictionaries methods. It signifies that the machine learning approach is still a trend for future research.

The literature review presented in this article has some limitations. Selection criteria are specific. This article does not include studies that are not following the research objectives. Moreover, the current study only investigates the last six-year publication (2016–2021). Despite its promising findings, the challenges to this research area need to be acknowledged. Educational contexts play an essential role in learning. Generalisability and reliability are still the main issues. The selected studies show that the experiments in generalisability of cross-domain context resulted in unsatisfactory performances. On the other hand, the reliability has potential to improve better in the future. In addition, most of the studies in this research area are in the method phase. However, there are still practical implementations in the topic, social, and summary measurements. In other measurements, there is a lack of evidence of the impact of the automated application in improving learning.

Future research will focus on tackling the challenges of automatic content analysis of asynchronous discussion forum transcripts. There is much room for improvement in this research area. The main challenge is to examine the automatic content analysis in different learning contexts, especially language and subject. Besides English, other languages need to be explored further. Furthermore, there is a need to develop automatic applications at the cognitive level. The practical implementation of this automatic content analysis approach in the online learning environment is expected to impact learners and teachers positively.

As for future work, we plan to develop an automated system in analysing the cognitive presence, social presence, and teaching presence of asynchronous discussion forums based on the CoI framework. The current study shows a lack of research in Indonesian language and blended learning research contexts. Hence, our future work will explore the discussion forum transcripts of blended learning in higher education. The context of the language used is Indonesian. Currently, online collaborative learning, especially content analysis, is developing in Indonesia (Junus et al., 2019, 2021; Purwandari et al., 2022; Tjhin et al., 2017). The Indonesian government also played a role by developing SPADA (spada.kemdikbud.go.id) as the national MOOCs platform. The role of asynchronous discussion has become significant and relevant in Indonesia because of its various development of IT infrastructures. The asynchronous approach enables learners who do not have adequate infrastructure to study anytime (Afify, 2019). Moreover, the study of automated content analysis of asynchronous discussion forums using the CoI framework has not yet begun in Indonesia.

The initial studies were conducted to prepare our Indonesian language datasets from STEM courses: linear algebra (Junus et al., 2019) and computer-aided instruction classes (Junus et al., 2021). These previous studies used manual content analysis to identify the social, cognitive, and teaching presences from discussion transcripts. The unit of analysis used was messages. Three experienced coders were involved in the coding process based on Community of Inquiry's coding protocol (Shea et al., 2010). The coding results of each coder were compared to obtain Fleiss's kappa as

interrater reliability coefficient. The content analysis results reported the level of agreements from moderate to almost perfect. Our future study will use the labelled datasets from these initial studies in developing an automated system to categorise cognitive, social, and teaching presences of Indonesian language discussion transcripts with machine learning and dictionary-based approaches.

Our research methodology consisted of three main phases: dataset preparation, model development, and implementation. Dataset generation is the first phase in developing the automated analysis system of Community of Inquiry. The purpose of this phase is to obtain a reliable and valid dataset. This phase consisted of five stages: planning, data collection, human-coding analysis, and dataset preparation. After getting the dataset, the next phase is model development. This phase aims to obtain the best model in classifying discussion messages into the Community of Inquiry categories using a machine-learning approach. Hence, there are two main stages: training and evaluation. Next, the best model obtained in the previous phase is implemented into an automated analysis system. This automated system should also be integrated where the online discussion takes place, in this case, in the Learning Management System (LMS). This methodology is used as our guidelines to develop the automated analysis of Community of Inquiry elements in online discussion forums. Figure 6 below shows our future research methodology.

6 Conclusions

This article aims to investigate the current state of the research related to analysing discussion transcripts automatically. The systematic literature review was conducted using the PRISMA framework and snowballing on five electronic academic and citation databases: IEEE Xplore, ACM Digital Library,
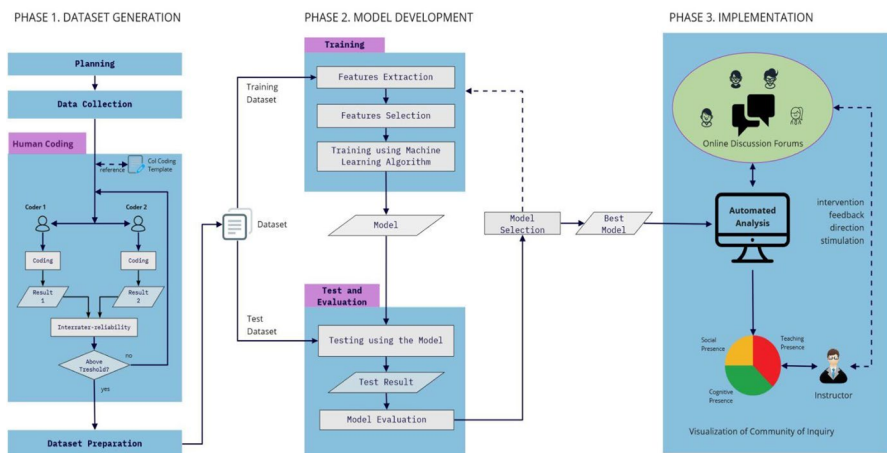
Fig. 6 Research methodology of future work. Source: Developed by the authors using online design tool Miro (https://miro.com/)

ScienceDirect, Scopus, and ERIC. The range of the years of publication of literature is between January 2016 and October 2021. Based on our eligibility criteria, 54 out of 414 articles are relevant to the automated approach in analysing discussion transcripts of online learning. Four in-depth aspects (i.e., measurement, learning context, methods, and theoretical frameworks used in these studies) were analyzed. This study found eight measurement dimensions: cognitive, social, relevance/interest, summary, pattern, behaviour, topic, and learning resources. Social, relevance, and cognitive were the top three measurements in the last six years. The CoI was the most widely used framework in analysing online learning discussion transcripts. It shows that the recent studies focus on helping teachers understand learning processes in the discussion forums. Seventy-two percent of studies investigated the data from fully online learning environments, especially MOOCs. Also, 78% of studies used English discussion transcripts as the dataset. In addition, 59% of studies used single-disciplinary datasets with STEM as the most-used knowledge area. In terms of method, machine learning was the most widely used method compared to other methods. Ninety-three percent of all studies used machine learning to analyse discussion transcripts. The experiments of studies show promising results in precision and accuracy. From multidisciplinary researches, the experiments show unsatisfactory results. Therefore, this research topic still has a big room for improvement in the future, especially in improving the reliability and generalisability of cross-domain context. Furthermore, the future direction of this research area should be related to the contexts of learning, such as language, learning environment, and subject.

Appendix

Table 16 A summary of selected studies in systematic literature review

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Almatrafi et al., 2018)	Classification	Urgent (post)	N/A	English	Fully online (MOOCs)	Machine Learning	•Adaboost •Logistic Regression •Random Forest	•weighted-F1 = 0.88 •F1 ≥ 0.70 •Kappa ≥ 0.63
(An et al., 2019)	Classification	Resource mention	N/A	English	Fully online (MOOCs)	Machine Learning	Incremental Neural Model based on LSTM–CRF	•Precision = 0.729 •Recall = 0.79 •F1 = 0.759
(Atapattu et al., 2016)	Clustering	Topic model	N/A	English	Fully online (MOOCs)	Machine Learning	•Latent Dirichlet Allocation (LDA) •Naïve Bayes	•F1 ≥ 0.69 •Macroaveraged F1 ≥ 0.896
(Barbosa et al., 2020)	Classification	Cognitive presence	Community of Inquiry (Garrison et al., 1999)	English and Portuguese	Fully online (Higher Education)	Machine Learning	Random Forest	•Accuracy = 67% •Cohen's kappa = 0.53

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Barbosa et al., 2021)	Classification	•Social presence •Cognitive presence	Community of Inquiry (Garrison et al., 1999)	English and Portuguese	Fully online (Higher Education)	Machine Learning Rule-based	•Random Forest •Automatic Text Translation	Social Presence (Affective) •Accuracy=91% •Cohen's kappa=0.49 Social Presence (Interactive) •Accuracy=96% •Cohen's kappa=0.92 Social Presence (Cohesive) •Accuracy=88% •Cohen's kappa=0.44 Cognitive Presence •Accuracy=83% •Cohen's kappa=0.69
(Boroujeni et al., 2017)	Classification	•Forum activity •Speech of acts (post) •Social network	Speech Acts (Arguello & Shaffer, 2015)	English	Fully online (MOOCs)	Machine Learning	•Random Forest •Iterative Classification •Social Network Analysis	•F1=0.77
(Bosch et al., 2020)	Classification	Private information	N/A	English	Online learning (Higher Education)	Machine Learning	•Dictionary-Based •Extra-Trees •Custom NN •Fusion	•Recall=0.954 •AUC=0.979

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Brinton et al., 2018)	Classification	Question or answer (post)	N/A	English	Fully online (MOOCs)	Machine Learning	Proposed Estimation Model	•Accuracy=86% •F1=0.65
	Clustering	Topic model				Machine Learning	Latent Dirichlet Allocation (LDA)	
(Chen et al., 2018)	Clustering	•Speaking behaviour •Social network	Speaking Behaviour (Wise et al., 2014)	English	Fully online (Higher Education)	Machine Learning	Lexical analysis	•Krippendorff's $\alpha > 0.60$
(Chen et al., 2021)	Clustering	Topic model	N/A	Chinese	Online learning (High School)	Machine Learning	Latent Dirichlet Allocation (LDA)	The students' performance of experimental group were better than the control group in socio-scientific reasoning. The experimental group used the proposed system (TAIFS-assisted online discussion)

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	*Learning Mode*	*Approach*	*Algorithm*	Result
(Cheng et al., 2021)	Classification	Cognitive Level	Bloom's Taxonomy (Anderson et al., 2001)	Chinese	Fully online (MOOCs)	Machine Learning	The proposed method combines bidirectional gated recurrent unit, convolutional neural network model, and attention mechanism (AM-BiGRU-CNN)	•Accuracy=84.21% •F1=0.9171
(Farrow et al., 2019)	Classification	Cognitive presence	Community of Inquiry (Garrison et al., 1999)	English	Fully online (Higher Education)	Machine Learning	Random Forest	•Accuracy=61.7% •Cohen's kappa=0.46 •Macro-averaged F1=0.58
(Feng et al., 2018)	Classification	Type of Thread	N/A	English, French, Chinese, and Spanish	Fully Online (MOOCs)	Machine Learning	Gradient Boosting Decision Tree (GBDT) with Linear Regression (LR)	•Accuracy≥78.1% •Average AUC=0.832
(Ferreira et al., 2020)	Classification Visualisation	•Social presence •Epistemic network	Community of Inquiry (Garrison et al., 1999)	English	Fully online (Higher Education)	Machine Learning	Classifier: Random Forest Epistemic Network Analysis	•Cohen's kappa≥0.49 •Accuracy≥81%

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Fu et al., 2018)	Clustering	•User groups •Sentiment	N/A	English	Fully online (MOOCs)	Machine Learning	Dense Subgraph Co-occurence	Usability study of the system functionality showed positive results
(Gašević et al., 2019)	Clustering	•Topic model •Social network •Epistemic network	N/A	English	Fully online (MOOCs)	Machine Learning	•Latent Dirichlet Allocation (LDA) •Social Network Analysis (SNA) •Epistemic Network Analysis (ENA)	Twelve topics were successfully identified by LDA to construct ENA
(Gottipati et al., 2019)	Clustering Summarisation	Topic model Summary	N/A	English	Online learning (Higher Education)	Machine Learning	K-mean Clustering TextRank Summariser and LSA Summariser	LSA summarization is slightly better than TextRank
(Hayati et al., 2019)	Classification	Cognitive presence	Community of Inquiry (Garrison et al., 1999)	English	N/A	Machine Learning	Naïve Bayes	•F1=0.7705 •Cohen's kappa=0.6866 •Accuracy=76.47% •Recall=0.689

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Hayati et al., 2020)	Classification	•Cognitive engagement •Cognitive presence	•ICAP (Chi & Wylie, 2014) •Community of Inquiry (Garrison et al., 1999)	N/A	Online learning	Machine Learning	Support Vector Machines (SVM)	•F1≥0.84 •Cohen's kappa≥0.79 •Accuracy≥85% •Precision=0.95 •Recall=0.93
(Hecking et al., 2016)	Classification	Speech of acts (post)	Speech Acts (Arguello & Shaffer, 2015)	English	Fully online (MOOCs)	Machine Learning	Random Forest and Iterative Classification	•F1≥0.59 •Precision≥0.61 •Recall≥0.53
(Huang et al., 2019)	Classification	Metacognition type (sentence)	N/A	English	Online learning (Higher Education)	Rule-based	Dictionary-based	•Spearman's rho≥0.859 •Cohen's kappa≥0.776
(Irish et al., 2020)	Clustering	Similarity (post)	N/A	English	Online learning (Higher Education)	Machine Learning	TF/IDF bag of words	Students' satisfaction of the tools: •Ease of use=56% •Ease of installation=66% •Ease of search=54% •Ability to find relevant posts=60%
(Joksimovic et al., 2020)	Classification	Speech acts	Speech Acts (Arguello & Shaffer, 2015)	English	Fully online (MOOCs)	Machine Learning	Hidden Markov Models (HMM) Latent Dirichlet Allocation (LDA)	HMM and LDA used to automatically identify speech act categories

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	*Learning Mode*	*Approach*	*Algorithm*	*Result*
(Kang et al., 2018)	Classification	• Sentiment • Social characteristics	N/A	Chinese	Online learning (Higher Education)	Rule-based	Dictionary-based Social Network Analysis	Strong positive correlation between the confusion sentiment and the overall scores of the students
(Kovanovic et al., 2017)	Classification	Cognitive presence	Community of Inquiry (Garrison et al., 1999)	English	Fully online (Higher Education)	Machine Learning	Random Forest	•Accuracy=70.3% •Cohen's kappa=0.63
(Le et al., 2018)	Classification	Critical (sentence)	N/A	English	Online learning (Higher Education)	Machine Learning	Multinomial Naïve Bayes	Accuracy=81.9%
(Liu et al., 2016)	Classification	Importance (post)	N/A	English	Fully online (MOOCs)	Machine Learning	Random Forest	Cohen's kappa=0.57
(Liu et al., 2018)	Clustering	•Topic model •Sentiment	N/A	Chinese	Fully online (Higher Education)	Rule-based	Temporal Emotion Topic Model (TETM) Lexicon-based	Experiments identified the representation of the topics (top ten words) based on each emotion (positive, negative, confused)
(Machado et al., 2019)	Classification	Relevance (post)	N/A	Portuguese	Online learning (Higher Education)	Machine Learning	•Rule-based approach •Multi-Layer Perceptron (MLP) •Support Vector Machine (SVM)	•Accuracy=84% (MLP and SVM) •Recall=0.80 (SVM) •Precision=0. 862 (MLP) •F1=0.8187 (SVM)

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method			Experiment
				Language	Learning Mode	Approach	Algorithm		Result
(Moore et al., 2019)	Classification	Cognitive processing	Community of Inquiry (Garrison et al., 1999)	English	Fully online (MOOCs)	Rule-based	Dictionary-based		Number of words in the dictionary was the strongest predictor for cognitive processing
(Moreno-Marcos et al., 2018)	Classification	Sentiment	N/A	English	Fully online (MOOCs)	Rule-based	Dictionary-based		• AUC=0.78 • Cohen's kappa=0.54
						Machine Learning	Random Forest		• AUC=0.85 • Cohen's kappa=0.58
(Moreno-Marcos et al., 2019)	Classification	Sentiment	N/A	English	Fully online (MOOCs)	Rule-based	Dictionary-based		The sentiments of discussion (good/fair/bad) were classified based on positive posts, negative posts, neutral posts, and percentage of positivity
(Neto et al., 2018)	Classification	Cognitive presence	Community of Inquiry (Garrison et al., 1999)	Portuguese	Fully online (Higher Education)	Machine Learning	Random Forest		• Accuracy=83% • Cohen's kappa=0.72
(Neto et al., 2021)	Classification	Cognitive presence	Community of Inquiry (Garrison et al., 1999)	Portuguese	Fully online (Higher Education)	Machine Learning	Random Forest		• Accuracy=76% • Cohen's kappa=0.55

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Ntourmas et al., 2019)	Classification	Relevance (starting post)	N/A	Greek	Fully online (MOOCs)	Machine Learning	Support Vector Machines (SVM)	•Accuracy=69% and 82% •Precision=0.70 and 0.84 •Recall=0.69 and 0.82 •F1=0.70 and 0.78
(Ntourmas et al., 2021)	Classification	Relevance (starting post)	N/A	Greek	Fully online (MOOCs)	Machine Learning	Weighted Decision Tree	•Accuracy$\geq$60% •$0.22\leq$Precision≤0.84 •$0.34\leq$Recall≤0.65 •$0.27\leq$F1≤0.72
(Peng et al., 2020)	Classification	Behaviour	N/A	Chinese	Fully online (Higher Education)	Rule-based	Dictionary-based	The proposed method (TI-EBTM) performed better in identifying higher-quality temporal topics than other models (Post-LDA and ATU-LDA)
	Clustering	Topic model				Machine Learning	Time Information-Emotion Behaviour Model (TI-EBTM): (extension of LDA)	

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Peng et al., 2021)	Classification	Behaviour	N/A	Chinese	Fully online (Higher Education)	Rule-based	Dictionary-based	The proposed method (SBTM) performed significantly better than the other topic models (SLDA, TLDA, L-SLDA, and D-TLDA)
	Clustering	Topic model				Machine Learning	Sentiment and Behaviour Topic Model (SBTM) (extension of LDA)	
(Rolim et al., 2019)	Classification	Student profile	N/A	English	Online learning (Higher Education)	Machine Learning	Support Vector Machines (SVM) and Genetic Algorithm	F1 = 98.76%
	Clustering	Topic model				Machine Learning	Latent Dirichlet Allocation (LDA)	Topic Extraction •Percentage of agreement 90% •Cohen's kappa=0,7936 Keyword Extraction •Percentage of agreement = 94.16% •Cohen kappa = 0,8727

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Setiawan et al., 2020)	Clustering	Topic model	N/A	Indonesian	Online learning (Higher Education)	Machine Learning	Probability Latent Semantic Analysis (PLSA)	•Precision=0.65 •Recall=0.50 •F1=0.57
(Sun et al., 2019)	Classification	Urgent (post)	N/A	English	Fully online (MOOCs)	Machine Learning	Recurrent Convolutional Neural Networks (RCNN)+LSTM	•Accuracy=84.8% •Precision=0.827 •Recall=0.760 •F1=0.786
(Wise & Cui, 2018a)	Classification	•Participation •Relevance (thread)	Post Classification (Stump et al., 2013)	English	Fully online (MOOCs)	Machine Learning	Binary L2 Regularised Logistic Regression and DIPTiC	Accuracy=88%
(Wise & Cui, 2018b)	Classification	•Relevance (thread) •Social network	Post Classification (Stump et al., 2013)	English	Fully online (MOOCs)	Machine Learning	Binary L2 Regularised Logistic Regression and DIPTiC	Accuracy=88%
(Wise et al., 2016)	Classification	Relevance (thread)	Post Classification (Stump et al., 2013)	English	Fully online (MOOCs)	Machine Learning	Binary L2 Regularised Logistic Regression	•Accuracy=0.80 •Cohen's kappa=0.61 •Recall=0.79 •Precision=0.79
(Wise et al., 2017a, 2017b)	Classification	•Relevance (thread) •Social network	Post Classification (Stump et al., 2013)	English	Fully online (MOOCs)	Machine Learning	Binary L2 Regularised Logistic Regression and DIPTiC	Accuracy=88%

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Wise et al., 2017a, 2017b)	Classification	Relevance (thread)	Post Classification (Stump et al., 2013)	English	Fully online (MOOCs)	Machine Learning	Binary L2 Regularised Logistic Regression	•Accuracy ˃ 0.80, •Kappa ˃ 0.52 •Precision ˃ 0.62 •Recall ˃ 0.72
(Wong & Li, 2016)	Association	Chance discovery	N/A	English	Blended	Rule-based	Co-occurrence analysis and Key graphs	The experiments showed the traceable pattern in chance discovery was the significant predictor for the academic performance of the students
(Wong et al., 2016)	Clustering	Topic model	N/A	English	N/A	Machine Learning	Latent Dirichlet Allocation (LDA)	The experiments showed the functionality of the proposed system (LDAvis): topic extraction and data visualization. However, there is no evaluation phase
(Wu & Wu, 2018)	Classification	•Cognitive behaviour •Cognitive gain	ICAP (Chi & Wylie, 2014)	English	Fully online (MOOCs)	Machine Learning	Logistic Regression	•Accuracy≥0.83 •AUC≥0.73
(Xie et al., 2018)	Classification	•Leader (post) •Social network	N/A	English	Fully online (Higher Education)	Machine Learning	Logistic Regression Model (LR)	•Accuracy=86% •Precision=0.912 •Recall=0.838 •F1=0.873

Table 16 (continued)

Studies	Category	Measurement	Theoretical Background (Framework / Model)	Dataset		Method		Experiment
				Language	Learning Mode	Approach	Algorithm	Result
(Yang et al., 2018)	Clustering	•Topic Model •Social learning network	N/A	English	Fully online (MOOCs)	Machine Learning	Latent Dirichlet Allocation (LDA)	•$0.71 \leq AUC \leq 0.97$ •$0.90 \leq ACC \leq 0.94$
(Yang et al., 2022)	Classification	Critical Thinking (Cognitive Presence)	Community of Inquiry (Garrison et al., 1999)	English	Fully online (Higher Education)	Machine Learning	Multi-Feature Fusion Modeling (MFFM) and Long Short-Term Memory–Convolutional Neural Network (LSTM-CNN)	•$Accuracy \geq 82.05\%$ •$0.5257 \leq Precision \leq 0.9131$ •$0.7472 \leq Recall \leq 0.8431$ •$0.6172 \leq F1 \leq 0.8767$
(Zarra et al., 2018)	Clustering	Topic model	N/A	English	Fully online (MOOCs)	Machine Learning	Latent Dirichlet Allocation (LDA)	•$Accuracy = 62.06\%$ •$Precision = 0.6206$ •$Recall = 0.7425$ •$F1 = 0.6761$
(Zou et al., 2021)	Classification	Social Presence	Community of Inquiry (Garrison et al., 1999)	English	Fully online (MOOCs)	Machine Learning	Bidirectional Encoder Representations from Transformers (BERT)	•$Accuracy = 83\%$ •$Precision = 0,82$ •$Recall = 0,80$ •$F1 = 0,81$

Declarations

References

Afify, M. K. (2019). The influence of group size in the asynchronous online discussions on the development of critical thinking skills, and on improving students' performance in online discussion forum. *International Journal of Emerging Technologies in Learning (IJET), 14*(5), 132–152. https://doi.org/10.3991/ijet.v14i05.9351

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics, 37*, 13–49. https://doi.org/10.1016/j.tele.2019.01.007

Almatrafi, O., Johri, A., & Rangwala, H. (2018). Needle in a haystack: Identifying learner posts that require urgent response in MOOC discussion forums. *Computers & Education, 118*, 1–9. https://doi.org/10.1016/j.compedu.2017.11.002

Altınay, Z. (2017). Evaluating peer learning and assessment in online collaborative learning environments. *Behaviour & Information Technology, 36*(3), 312–320. https://doi.org/10.1080/0144929X.2016.1232752

An, Y.-H., Pan, L., Kan, M.-Y., Dong, Q., & Fu, Y. (2019). Resource mention extraction for mooc discussion forums. *IEEE Access, 7*, 87887–87900. https://doi.org/10.1109/ACCESS.2019.2924250

Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (2001). *Taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Arguello, J., & Shaffer, K. (2015). Predicting speech acts in MOOC forum posts. *Proceedings of the International AAAI Conference on Web and Social Media*, *9*(1), 2–11. https://ojs.aaai.org/index.php/ICWSM/article/view/14604

Atapattu, T., Falkner, K., & Tarmazdi, H. (2016). Topic-wise classification of MOOC discussions: A visual analytics approach. *Proceedings of International Conference on Educational Data Mining, 9*, 276–281.

Baker, R. S. (2019). Challenges for the future of educational data mining: The Baker learning analytics prizes. *Journal of Educational Data Mining, 11*(1), 1–17. https://doi.org/10.5281/ZENODO.3554745

Barbosa, G., Camelo, R., Cavalcanti, A. P., Miranda, P., Mello, R. F., Kovanović, V., & Gašević, D. (2020). Towards automatic cross-language classification of cognitive presence in online discussions. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, 605–614,*. https://doi.org/10.1145/3375462.3375496

Barbosa, A., Ferreira, M., Ferreira Mello, R., Dueire Lins, R., & Gasevic, D. (2021). The impact of automatic text translation on classification of online discussions for social and cognitive presences. *LAK21: 11th International Learning Analytics and Knowledge Conference*, 77–87. https://doi.org/10.1145/3448139.3448147

Bates, A. W. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning*. Tony Bates Associates Ltd. https://opentextbc.ca/teachinginadigitalage/

Beckmann, J., & Weber, P. (2016). Cognitive presence in virtual collaborative learning: Assessing and improving critical thinking in online discussion forums. *Interactive Technology and Smart Education*, *13*(1), 52–70. https://www.learntechlib.org/p/193097/

Berry, M. W. (Ed.). (2004). *Survey of text mining: Clustering, classification, and retrieval*. Springer.

Boroujeni, M. S., Hecking, T., Hoppe, H. U., & Dillenbourg, P. (2017). Dynamics of MOOC discussion forums. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference, 128–137*. https://doi.org/10.1145/3027385.3027391

Bosch, N., Crues, R. W., Shaik, N., & Paquette, L. (2020). Hello,[REDACTED]: Protecting Student Privacy in Analyses of Online Discussion Forums. *Proceedings of The 13th International Conference on Educational Data Mining*, pp 39–49.

Brinton, C. G., Buccapatnam, S., Zheng, L., Cao, D., Lan, A. S., Wong, F. M. F., Ha, S., Chiang, M., & Poor, H. V. (2018). On the efficiency of online social learning networks. *IEEE/ACM Transactions on Networking, 26*(5), 2076–2089. https://doi.org/10.1109/TNET.2018.2859325

Brookfield, S. (1995). Adult learning: An overview. *International Encyclopedia of Education, 10*, 375–380.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *ArXiv:2005.14165[Cs]*. http://arxiv.org/abs/2005.14165

Chang, B. (2019). Reflection in learning. *Online Learning*, *23*(1), 95–110. https://doi.org/10.24059/olj.v23i1.1447

Chen, B., Chang, Y.-H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education, 37*, 21–30. https://doi.org/10.1016/j.iheduc.2017.12.002

Chen, C.-M., Li, M.-C., Chang, W.-C., & Chen, X.-X. (2021). Developing a Topic Analysis Instant Feedback System to facilitate asynchronous online discussion effectiveness. *Computers & Education, 163*, 104095. https://doi.org/10.1016/j.compedu.2020.104095

Cheng, Y., Cai, Y., Chen, H., Cai, Z., Wu, G., & Huang, J. (2021). A cognitive level evaluation method based on a deep neural network for online learning: from a bloom's taxonomy of cognition objectives perspective. *Frontiers in psychology, 12*, 661235. https://doi.org/10.3389/fpsyg.2021.661235

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *ArXiv:1810.04805[Cs]*. http://arxiv.org/abs/1810.04805

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook for research for educational communications and technology* (pp. 170–198). Simon & Schuster Macmillan.

Fahy, P. J. (2001). Addressing some common problems in transcript analysis. *The International Review of Research in Open and Distributed Learning*, *1*(2), 133–138. https://doi.org/10.19173/irrodl.v1i2.321

Farrow, E., Moore, J., & Gašević, D. (2019). Analysing discussion forum data: A replication study avoiding data contamination. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 170–179. https://doi.org/10.1145/3303772.3303779

Feng, L., Lu, H., Liu, S., Liu, G., & Luo, S. (2018). Automatic feature learning for mooc forum thread classification. *Proceedings of the 2018 International Conference on Big Data and Computing*, 65–70. https://doi.org/10.1145/3220199.3220201

Ferreira, M., Rolim, V., Mello, R. F., Lins, R. D., Chen, G., & Gašević, D. (2020). Towards automatic content analysis of social presence in transcripts of online discussions. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, 141–150,*. https://doi.org/10.1145/3375462.3375495

Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9*(6), e1332. https://doi.org/10.1002/widm.1332

Fu, S., Wang, Y., Yang, Y., Bi, Q., Guo, F., & Qu, H. (2018). Visforum: A visual analysis system for exploring user groups in online forums. *ACM Transactions on Interactive Intelligent Systems (TiiS), 8*(1), 1–21. https://doi.org/10.1145/3162075

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education, 9*(1), 1–8. https://doi.org/10.1016/j.iheduc.2005.11.001

Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior, 92*, 562–577. https://doi.org/10.1016/j.chb.2018.07.003

Gonzalez, H. B., & Kuenzi, J. J. (2012). *Science, technology, engineering, and mathematics (STEM) education: A primer*. Congressional Research Service, Library of Congress. https://sgp.fas.org/crs/misc/R42642.pdf

Gottipati, S., Shankararaman, V., & Ramesh, R. (2019). TopicSummary: A tool for analyzing class discussion forums using topic based summarizations. *IEEE Frontiers in Education Conference (FIE), 2019*, 1–9. https://doi.org/10.1109/FIE43999.2019.9028526

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 397–431. https://doi.org/10.2190/7MQV-X9UJ-C7Q3-NRAG

Harasim, L. (2017). *Learning Theory and Online Technologies (2nd ed.).* Routledge. https://doi.org/10.4324/9781315716831

Hayati, H., Chanaa, A., KhalidiIdrissi, M., & Bennani, S. (2019). Doc2vec & naïve bayes: Learners' cognitive presence assessment through asynchronous online discussion tq transcripts. *International Journal of Emerging Technologies in Learning (IJET), 14*(8), 70. https://doi.org/10.3991/ijet.v14i08.9964

Hayati, H., Khalidi Idrissi, M., & Bennani, S. (2020). Automatic classification for cognitive engagement in online discussion forums: Text mining and machine learning approach. *Proceedings of International Conference Artificial Intelligence in Education (AIED), 21*(2), 114–118. https://doi.org/10.1007/978-3-030-52240-7_21

Hecking, T., Chounta, I.-A., & Hoppe, H. U. (2016). Investigating social and semantic user roles in MOOC discussion forums. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 198–207. https://doi.org/10.1145/2883851.2883924

Henri, F. (1992). Computer conferencing and content analysis. In A. R. Kaye (Ed.), *Collaborative Learning Through Computer Conferencing* (pp. 117–136). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-77684-7_8

Huang, E., Valdiviejas, H., & Bosch, N. (2019). I'm sure! Automatic detection of metacognition in online course discussion forums. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1–7. https://doi.org/10.1109/ACII.2019.8925506

Irish, I., Finkelberg, R., Nkemelu, D., Gujrania, S., Padiyath, A., Raman, S., Tailor, C., Arriaga, R., & Starner, T. (2020). Parqr: Automatic post suggestion in the piazza online forum to support degree seeking online masters students. *Proceedings of the Seventh ACM Conference on Learning @ Scale*, 125–134. https://doi.org/10.1145/3386527.3405914

Jayakumar, N., Brunckhorst, O., Dasgupta, P., Khan, M. S., & Ahmed, K. (2015). E-learning in surgical education: A systematic review. *Journal of Surgical Education, 72*(6), 1145–1157. https://doi.org/10.1016/j.jsurg.2015.05.008

Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. Sage Publications.

Joksimovic, S., Jovanovic, J., Kovanovic, V., Gasevic, D., Milikic, N., Zouaq, A., & van Staalduinen, J. P. (2020). Comprehensive analysis of discussion forum participation: From speech acts to discussion dynamics and course outcomes. *IEEE Transactions on Learning Technologies, 13*(1), 38–51. https://doi.org/10.1109/TLT.2019.2916808

Junus, K., Santoso, H. B., & Ahmad, M. (2021). Experiencing the community of inquiry framework using asynchronous online role-playing in computer-aided instruction class. *Education and Information Technologies, 1–27,*. https://doi.org/10.1007/s10639-021-10670-5

Junus, K., Suhartanto, H., R-Suradijono, S. H., Santoso, H. B., & Sadita, L. (2019). The community of inquiry model training using the cognitive apprenticeship approach to improve students' learning strategy in the asynchronous discussion forum. *Journal of Educators Online*, *16*(1), 69–86.

Kang, L., Liu, Z., Su, Z., Li, Q., & Liu, S. (2018). Analyzing the relationship among learners' social characteristics, sentiments in a course forum and learning outcomes. *Seventh International Conference of Educational Innovation through Technology (EITT), 2018*, 210–213. https://doi.org/10.1109/EITT.2018.00049

Koedinger, K. R., D'Mello, S., McLaughlin, E. A., Pardos, Z. A., & Rosé, C. P. (2015). Data mining and education. *Wires Cognitive Science, 6*(4), 333–353. https://doi.org/10.1002/wcs.1350

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., & Siemens, G. (2016). Towards automated content analysis of discussion transcripts: A cognitive presence case. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 15–24,*. https://doi.org/10.1145/2883851.2883950

Kovanovic, V., Joksimovic, S., Gašević, D., Hatala, M., and Siemens, G. (2017). Content Analytics: The Definition, Scope, and an Overview of Published Research. In C. Lang, G. Siemens, A.F. Wise, and D. Gašević (Eds.), *The Handbook of Learning Analytics* (pp. 77–92). Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17.007

Kowsari, K., JafariMeimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. *Information, 10*(4), 150. https://doi.org/10.3390/info10040150

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (Fourth Edition). Sage Publications.

Ku, H. Y., Tseng, H. W., & Akarasriworn, C. (2013). Collaboration factors, teamwork satisfaction, and student attitudes toward online collaborative learning. *Computers in Human Behavior, 29*(3), 922–929. https://doi.org/10.1016/j.chb.2012.12.019

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Le, K., Tarmazdi, H., Vivian, R., Falkner, K., Szabo, C., & Falkner, N. (2018). Directing teacher focus in computer science online learning environments. *International Conference on Learning and Teaching in Computing and Engineering (LaTICE), 2018*, 119–125. https://doi.org/10.1109/LaTICE.2018.00014

Lee, K. (2007). Online collaborative case study learning. *Journal of College Reading and Learning, 37*(2), 82–100. https://doi.org/10.1080/10790195.2007.10850199

Liu, Z., Rudian, S., Yang, C., Sun, J., & Liu, S. (2018). Tracking the dynamics of SPOC discussion forums: A temporal emotion-topic modeling approach. *Seventh International Conference of Educational Innovation through Technology (EITT), 2018*, 174–179. https://doi.org/10.1109/EITT.2018.00042

Liu, W., Kidziński, Ł., & Dillenbourg, P. (2016). Semiautomatic annotation of mooc forum posts. In Y. Li, M. Chang, M. Kravcik, E. Popescu, R. Huang, Kinshuk, & N.-S. Chen (Eds.), *State-of-the-Art and Future Directions of Smart Learning* (pp. 399–408). Springer Singapore. https://doi.org/10.1007/978-981-287-868-7_48

Machado, C. J. R., Maciel, A. M. A., Rodrigues, R. L., & Menezes, R. (2019). An approach for thematic relevance analysis applied to textual contributions in discussion forums. *International Journal of Distance Education Technologies (IJDET), 17*(3), 37–51. https://doi.org/10.4018/IJDET.2019070103

Mohammed, A. H. K., Jebamikyous, H. H., Nawara, D., & Kashef, R. (2021). IoT text analytics in smart education and beyond. *Journal of Computing in Higher Education, 33*(3), 779–806. https://doi.org/10.1007/s12528-021-09295-x

Moon, J., & Park, Y. (2021). A scoping review on open educational resources to support interactions of learners with disabilities. *The International Review of Research in Open and Distributed Learning*, *22*(2), 314–341. https://doi.org/10.19173/irrodl.v22i1.5110

Moore, R. L., Oliver, K. M., & Wang, C. (2019). Setting the pace: Examining cognitive processing in MOOC discussion forums with automatic text analysis. *Interactive Learning Environments, 27*(5–6), 655–669. https://doi.org/10.1080/10494820.2019.1610453

Morais, Y., & Sampson, S. (2010). A content analysis of chat transcripts in the Georgetown law library. *Legal Reference Services Quarterly, 29*(3), 165–178. https://doi.org/10.1080/02703191003751289

Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Estévez-Ayres, I., & Kloos, C. D. (2018). Sentiment analysis in MOOCs: A case study. *IEEE Global Engineering Education Conference (EDUCON), 2018*, 1489–1496. https://doi.org/10.1109/EDUCON.2018.8363409

Moreno-Marcos, P. M., Alario-Hoyos, C., Munoz-Merino, P. J., Estevez-Ayres, I., & Kloos, C. D. (2019). A learning analytics methodology for understanding social interactions in MOOCs. *IEEE Transactions on Learning Technologies, 12*(4), 442–455. https://doi.org/10.1109/TLT.2018.2883419

Neto, V., Rolim, V., Pinheiro, A., Lins, R. D., Gasevic, D., & Mello, R. F. (2021). Automatic content analysis of online discussions for cognitive presence: A study of the generalizability across educational contexts. *IEEE Transactions on Learning Technologies, 14*(3), 299–312. https://doi.org/10.1109/TLT.2021.3083178

Neto, V., Rolim, V., Ferreira, R., Kovanović, V., Gašević, D., Dueire Lins, R., & Lins, R. (2018). Automated analysis of cognitive presence in online discussions written in portuguese. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *Lifelong Technology-Enhanced Learning* (Vol. 11082, pp. 245–261). Springer International Publishing. https://doi.org/10.1007/978-3-319-98572-5_19

Neuendorf, K. A. (2002). *The content analysis guidebook*. Sage Publications.

Newman, D. R. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology Journal, 3*(2), 56–77.

Ntourmas, A., Avouris, N., Daskalaki, S., & Dimitriadis, Y. (2019). Comparative study of two different MOOC forums posts classifiers: analysis and generalizability issues. *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, 1–8. https://doi.org/10.1109/IISA.2019.8900682

Ntourmas, A., Dimitriadis, Y., Daskalaki, S., & Avouris, N. (2021). Classification of discussions in MOOC forums: an incremental modeling approach. *Proceedings of the Eighth ACM Conference on Learning @ Scale*, 183–194. https://doi.org/10.1145/3430895.3460137

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ (Clinical Research Ed), 372*, n160. https://doi.org/10.1136/bmj.n160

Palloff, R. M., & Pratt, K. (2010). *Collaborating online: Learning together in community* (Vol. 32). John Wiley & Sons.

Peixoto, B., Pinto, R., Melo, M., Cabral, L., & Bessa, M. (2021). Immersive virtual reality for foreign language education: A PRISMA systematic review. *IEEE Access, 9*, 48952–48962. https://doi.org/10.1109/ACCESS.2021.3068858

Peng, X., Han, C., Ouyang, F., & Liu, Z. (2020). Topic tracking model for analyzing student-generated posts in SPOC discussion forums. *International Journal of Educational Technology in Higher Education, 17*(1), 35. https://doi.org/10.1186/s41239-020-00211-4

Peng, X., Xu, Q., & Gan, W. (2021). SBTM: A joint sentiment and behaviour topic model for online course discussion forums. *Journal of Information Science, 47*(4), 517–532. https://doi.org/10.1177/0165551520917120

Poldner, E., Simons, P. R. J., Wijngaards, G., & van der Schaaf, M. F. (2012). Quantitative content analysis procedures to analyse students' reflective essays: A methodological review of psychometric and edumetric aspects. *Educational Research Review, 7*(1), 19–37. https://doi.org/10.1016/j.edurev.2011.11.002

Purwandari, E. P., Junus, K., & Santoso, H. B. (2022). Exploring e-learning community of inquiry framework for engineering education. *International Journal of Instruction, 15*(1), 619–632.

Rolim, V., Ferreira Leite de Mello, R., Ferreira, M., Pinheiro Cavalcanti, A., & Lima, R. (2019). Identifying students' weaknesses and strengths based on online discussion using topic modeling.

2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), 2161–
377X, 63–65. https://doi.org/10.1109/ICALT.2019.00020

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated sur-
vey. *WIREs Data Mining and Knowledge Discovery, 10*(3), e1355. https://doi.org/10.1002/
widm.1355

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological issues in the content
analysis of computer conference transcripts. *International Journal of Artificial Intelligence in
Education (IJAIED), 12*(1), 8–22.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In
K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 97–118). Cambridge Uni-
versity Press.

Scharkow, M. (2017). Content analysis, automatic. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.),
The International Encyclopedia of Communication Research Methods (1st ed., pp. 1–14).
Wiley. https://doi.org/10.1002/9781118901731.iecrm0043

Setiawan, R., Budiharto, W., Kartowisastro, I. H., & Prabowo, H. (2020). Finding model through
latent semantic approach to reveal the topic of discussion in discussion forum. *Education and
Information Technologies, 25*(1), 31–50. https://doi.org/10.1007/s10639-019-09901-7

Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., Valchova, A., & Rangan, P.
(2010). A re-examination of the community of inquiry framework: Social network and content
analysis. *The Internet and Higher Education, 13*(1), 10–21. https://doi.org/10.1016/j.iheduc.
2009.11.002

Strijbos, J.-W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are
they talking about? *Computers & Education, 46*(1), 29–48. https://doi.org/10.1016/j.compedu.
2005.04.002

Stump, G. S., DeBoer, J., Whittinghill, J., & Breslow, L. (2013). Development of a framework to
classify MOOC discussion forum posts: Methodology and challenges. *NIPS Workshop on Data
Driven Education*. 1–20.

Sun, X., Guo, S., Gao, Y., Zhang, J., Xiao, X., & Feng, J. (2019). Identification of urgent posts in
MOOC discussion forums using an improved RCNN. *2019 IEEE World Conference on Engi-
neering Education (EDUNINE)*, 1–5.https://doi.org/10.1109/EDUNINE.2019.8875845

Thompson, L., & Ku, H.-Y. (2006). A case study of online collaborative learning. *Quarterly Review of
Distance Education, 7*(4), 361–375.

Tjhin, V. U., Rahayu, A., & Soraya, K. (2017). Evaluating the performance of students through col-
laborative learning: Case study: Distance education program in Indonesia. *2017 10th Interna-
tional Conference on Human System Interactions (HSI)*, 98–103. https://doi.org/10.1109/HSI.
2017.8005006

Turner, G., & Brass, K. (2014). *Mapping the humanities, arts and social sciences in Australia*. Aus-
tralian Academy of the Humanities.

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bib-
liometric mapping. *Scientometrics, 84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Wise, A. F., & Cui, Y. (2018a). Learning communities in the crowd: Characteristics of content related
interactions and social relationships in MOOC discussion forums. *Computers & Education,
122*, 221–242. https://doi.org/10.1016/j.compedu.2018.03.021

Wise, A. F., Hausknecht, S. N., & Zhao, Y. (2014). Attending to others' posts in asynchronous dis-
cussions: Learners' online "listening" and its relationship to speaking. *International Jour-
nal of Computer-Supported Collaborative Learning, 9*(2), 185–209. https://doi.org/10.1007/
s11412-014-9192-9

Wise, A. F., Cui, Y., & Vytasek, J. (2016). Bringing order to chaos in MOOC discussion forums
with content-related thread identification. *Proceedings of the Sixth International Conference on
Learning Analytics & Knowledge, 188–197*,. https://doi.org/10.1145/2883851.2883916

Wise, A. F., Cui, Y., Jin, W., & Vytasek, J. (2017b). Mining for gold: Identifying content-related
MOOC discussion threads across domains through linguistic modeling. *The Internet and
Higher Education, 32*, 11–28. https://doi.org/10.1016/j.iheduc.2016.08.001

Wise, A. F., & Cui, Y. (2018b). Unpacking the relationship between discussion forum participation
and learning in MOOCs: Content is key. *Proceedings of the 8th International Conference on
Learning Analytics and Knowledge*, 330–339. https://doi.org/10.1145/3170358.3170403

Wise, A. F., Cui, Y., & Jin, W. Q. (2017a). Honing in on social learning networks in MOOC forums:
Examining critical network definition decisions. *Proceedings of the Seventh International*

Learning Analytics & Knowledge Conference, 383–392.https://doi.org/10.1145/3027385.3027446

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, 1–10. https://doi.org/10.1145/2601248.2601268

Wong, G. K. W., & Li, S. Y. K. (2016). Academic performance prediction using chance discovery from online discussion forums. *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, *1*, 706–711. https://doi.org/10.1109/COMPSAC.2016.44

Wong, G. K. W., Li, S. Y. K., & Wong, E. W. Y. (2016). Analyzing academic discussion forum data with topic detection and data visualization. *2016 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 109–115. https://doi.org/10.1109/TALE.2016.7851779

Wu, Y., & Wu, W. (2018). A learning analytics system for cognition analysis in online learning community. In L. H. U & H. Xie (Eds.), *Web and Big Data* (Vol. 11268, pp. 243–258). Springer International Publishing. https://doi.org/10.1007/978-3-030-01298-4_21

Xie, K., Di Tosto, G., Lu, L., & Cho, Y. S. (2018). Detecting leadership in peer-moderated online collaborative learning through text mining and social network analysis. *The Internet and Higher Education, 38*, 9–17. https://doi.org/10.1016/j.iheduc.2018.04.002

Yang, J., Du, X., Hung, J.-L., & Tu, C. (2022). Analyzing online discussion data for understanding the student's critical thinking. *Data Technologies and Applications, 56*(2), 303–326. https://doi.org/10.1108/DTA-04-2021-0088

Yang, T.-Y., Brinton, C. G., & Joe-Wong, C. (2018). Predicting learner interactions in social learning networks. *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 1322–1330. https://doi.org/10.1109/INFOCOM.2018.8485927

Zarra, T., Chiheb, R., Faizi, R., & El Afia, A. (2018). Student interactions in online discussion forums: Visual analysis with lda topic models. *Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications, 1–5*,. https://doi.org/10.1145/3230905.3230920

Zou, W., Hu, X., Pan, Z., Li, C., Cai, Y., & Liu, M. (2021). Exploring the relationship between social presence and learners' prestige in MOOC discussion forums using automated content analysis and social network analysis. *Computers in Human Behavior, 115*, 106582. https://doi.org/10.1016/j.chb.2020.106582

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.