Before and during COVID-19: A Cohesion Network Analysis of students' online participation in moodle courses

Before and during COVID-19: A Cohesion Network Analysis of students' online participation in

moodle courses

Maria-Dorinela Dascalu[1], Stefan Ruseti[1], Mihai Dascalu[1,2], Danielle S. McNamara[3],

Mihai Carabas[1], Traian Rebedea[1], and Stefan Trausan-Matu[1,2]

[1] University Politehnica of Bucharest

[2] Academy of Romanian Scientists

[3] Arizona State University

Before and during COVID-19: A Cohesion Network Analysis of students' online participation in moodle courses

Abstract

The COVID-19 pandemic has changed the entire world, while the impact and usage of online learning environments has greatly increased. This paper presents a new version of the ReaderBench framework, grounded in Cohesion Network Analysis, which can be used to evaluate the online activity of students as a plug-in feature to Moodle. A Recurrent Neural Network with LSTM cells that combines global features, including participation and initiation indices, with a time series analysis on timeframes is used to predict student grades, while multiple sociograms are generated to observe interaction patterns. Students' behaviors and interactions are compared before and during COVID-19 using two consecutive yearly instances of an undergraduate course in Algorithm Design, conducted in Romanian using Moodle. The COVID-19 outbreak generated an off-balance, a drastic increase in participation, followed by a decrease towards the end of the semester, compared to the academic year 2018-2019 when lower fluctuations in participation were observed. The prediction model for the 2018-2019 academic year is partially generalizable to the second year, but explains a considerably lower variance (R^2 = 0.13). In addition to the quantitative analysis, a qualitative analysis of changes in student behaviors using comparative sociograms further supported conclusions that there were drastic changes in student behaviors observed as a function of the COVID-19 pandemic.

Before and during COVID-19: A Cohesion Network Analysis of students' online participation in moodle courses

Maria-Dorinela Dascalu [a], Stefan Ruseti [a], Mihai Dascalu [a,b,*], Danielle S. McNamara [c], Mihai Carabas [a], Traian Rebedea [a], Stefan Trausan-Matu [a,b]

[a] *Faculty of Automatic Control and Computers, University Politehnica of Bucharest, Bucharest, Romania*
[b] *Academy of Romanian Scientists, Bucharest, Romania*
[c] *Arizona State University, Department of Psychology, Arizona, USA*

ARTICLE INFO

ABSTRACT

The COVID-19 pandemic has changed the entire world, while the impact and usage of online learning environments has greatly increased. This paper presents a new version of the *ReaderBench framework,* grounded in Cohesion Network Analysis, which can be used to evaluate the online activity of students as a plug-in feature to Moodle. A Recurrent Neural Network with LSTM cells that combines global features, including participation and initiation indices, with a time series analysis on timeframes is used to predict student grades, while multiple sociograms are generated to observe interaction patterns. Students' behaviors and interactions are compared before and during COVID-19 using two consecutive yearly instances of an undergraduate course in Algorithm Design, conducted in Romanian using Moodle. The COVID-19 outbreak generated an off-balance, a drastic increase in participation, followed by a decrease towards the end of the semester, compared to the academic year 2018–2019 when lower fluctuations in participation were observed. The prediction model for the 2018–2019 academic year obtained an R^2 of 0.27, while the model for the second year obtained a better R^2 of 0.34, a value arguably attributable to an increased volume of online activity. Moreover, the best model from the first academic year is partially generalizable to the second year, but explains a considerably lower variance ($R^2 = 0.13$). In addition to the quantitative analysis, a qualitative analysis of changes in student behaviors using comparative sociograms further supported conclusions that there were drastic changes in student behaviors observed as a function of the COVID-19 pandemic.

1. Introduction

The COVID-19 pandemic has changed the entire world, from businesses, economy, to learning institutions. One of the most important challenges has been how to continue providing high quality education to students while confronted with severe restrictions in face-to-face contact. Prior to this global crisis, increasing numbers of universities and schools were already moving toward providing learners with some forms of online learning environments (OLEs; Crawford et al., 2020). The onset of COVID-19 has moved online education from being an add-on or alternative to a necessity. Many, if not most educators have been forced to transform their face-to-face learning activities and adapt to online learning modalities. Consequently, OLEs have become indispensable for both students and teachers during the COVID-19 pandemic.

Online education, when the internet works, has multiple benefits.

For example, OLEs can help facilitate access to resources and automated assessment, and afford means to share opinions and discuss various issues beyond geographical boundaries. Various online learning platforms are currently available to instructors and students, such as: Course Management Systems (CMS), Massive Open Online Course (MOOC), and Small Private Online Course (SPOC). Course Management Systems allow tutors to provide course content, while encouraging learning and collaboration between learners. Many CMSs have been developed over the years, including: a) commercial platforms, such as Blackboard (www.blackboard.com), Google Classroom (https://classroom.google.com), and Pearson Online Learning Services (https://www.pearson.com); and b) open source alternatives (Pappas, 2019), including Moodle (http://moodle.org), Edmodo (https://www.edmodo.com), Chamilo (https://chamilo.org/en/), Open EDX (https://open.edx.org), and Totara Learn (https://www.totaralearning.com).

Regardless of the platform, evaluating students in online environments can present an extremely time-consuming and challenging task. Instructors lack the face-to-face time that affords various means of qualitative assessment and individualized feedback, and they are confronted with multiple sources of data, which at times can seem arbitrary, overwhelming and difficult to explore. Our objective here is three-fold. First, we describe a new version of the *ReaderBench* framework (M. Dascalu, Dessus, Bianco, Trausan-Matu, & Nardy, 2014; M. Dascalu, McNamara, Trausan-Matu, & Allen, 2018) that can be used to evaluate online activity of students as a plug-in feature to Moodle. The new functionalities target the analysis of students' behaviors, the modeling of their interactions, as well as predicting student grades based on their online participation. Second, we evaluate the utility of this tool in supporting teachers by predicting students' course grades based on their current activity. Finally, we compare students' behaviors and interactions before and during COVID-19 in terms of differences derived from our wide range of indices, while also considering comparative interactive views illustrating their interactions, and providing a qualitative analysis from the tutor's perspective.

1.1. Current study objectives

The subject of this study is an undergraduate course on Algorithm Design offered at University Politehnica of Bucharest, which moved entirely online due to the COVID-19 pandemic. Courses, laboratories, and study groups transitioned to Microsoft Teams using online streaming, while course resources, mandatory weekly tests, discussions on forums, and homework, shifted to Moodle. Microsoft Teams was newly introduced in 2020 due to the COVID-19 pandemic, whereas Moodle had been used over the past decade. As such, student behaviors, interactions, and course performance were available for both 2018–2019 and 2019–2020 academic years, each reflecting different external conditions and demands, before versus during the COVID-19 pandemic.

Our overarching objective is to evaluate students' behaviors and interaction patterns and predict student grades based on their activity in online forum discussions and click-stream data. In *ReaderBench*, Cohesion Network Analysis (CNA; M. Dascalu, Trausan-Matu, McNamara, & Dessus, 2015) is used to assess semantic cohesion among students' posts. Weekly CNA sociograms are generated to examine how the interactions between peers and with tutors evolve from one week to the next (see e. g., Sirbu, Dascalu, Crossley, McNamara, & Trausan-Matu, 2019; Sirbu et al., 2018). These visualizations also provide insights into students' behaviors in association with course events, such as deadlines, assignments, tests, and exams. The visualizations are designed for teachers to follow the evolution of students in term of interactions, interactivity, and online participation, enabling them to intervene when they notice a decrease in participation or inactivity, thus increasing students' chances of passing or obtaining a better grade. Various sources of information from CNA and students' behaviors (e.g., from clickstream and log data) are combined to predict student grades. We also provide a qualitative analysis of the CNA visualizations based on observations of one of the authors, who has over 25 years of experience in teaching the Algorithm Design course, combined with extensive experience in conducting research on topics related to Computer-Supported Collaborative Learning (CSCL) and Intelligent Tutoring Systems (ITS). In contrast to the previous studies performed by M.-D. Dascalu et al. (2020), M. Dascalu, McNamara, et al. (2018) and Crossley, Paquette, Dascalu, McNamara, and Baker (2016), the entire processing flow is integrated and performed in Python. In this current version, we introduce an integrated pipeline that accounts for all types of indices (CNA, time series, and textual complexity), and data derived from clickstream logs are also integrated in the final predictions; although used separately in previous analyses, participation and initiation indices are combined for the first time. Moreover, a neural network for predicting course grades was introduced, while the visualizations were updated.

The COVID-19 pandemic has created new challenges that present some increased opportunities in terms of a wider adoption of Online Learning Environments (OLEs) at all educational levels, starting from elementary schools to universities (Crawford et al., 2020). In contrast to the wide range of studies centered on quantifying the impact of COVID-19 in education (Cao et al., 2020) or on using OLEs to maintain the level of education (Chick et al., 2020), our aim is to enhance the utility of OLEs for students and instructors. We introduce enhanced instruments to automatically assess student engagement in Moodle and observe different behavioral patterns. We do so by using state of the art Natural Language Processing techniques. Our approach builds on previous methods in which CNA is applied to forum discussion threads and it is combined with textual complexity indices reflective of students' writing style, and longitudinal analyses applied to click-stream data. Finally, we predict student performance using a recurrent neural network model that considers time series analyses on timeframes, in addition to the CNA, textual complexity, and longitudinal analysis indices.

2. Background literature

2.1. Online learning environments (OLEs)

OLEs are increasingly used by people around the world because they facilitate quick access to resources and information, allow users to share opinions and ideas, and even engage in open debates. Learners share their experiences and opinions and search for answers in online environments, while tutors and instructors share their knowledge and expertise. Moreover, in addition to the facilities brought to learners and instructors, OLEs have opened up new research areas, such as modeling members' participation, analyzing interactions, and identifying particular interaction patterns within the community (Moore, Dickson-Deane, & Galyen, 2011; Tu, 2002; Weidlich & Bastiaens, 2019).

Online learning is constantly evolving and has changed considerably since its first appearance. In 1989, the University of Phoenix, one of the pioneers in online education, offered the first online program which became the school's main focus (https://www.britannica.com/topic/University-of-Phoenix). Ten years later, the first entirely web-based university – Jones University – became accredited and, only one year later, the University of Texas provided a number of online classes on a website that included quizzes, surveys, grades, and calendars. Thus, a new industry emerged – online education technology. Dave Cormier from the University of Prince Edward Island coined the term Massive Open Online Course (MOOC) in 2008. Various MOOC environments appeared in subsequent years, including Coursera, edX, and Udacity (Achieve Virtual, n.d.). Online learning improved distance learning and began to act as a substitute for face-to-face classes. As technology continues to evolve at a rapid pace, institutions striving to provide high quality education are required to use the best alternatives (Hiltz & Turoff, 2005).

Engaging learners is a complex task for instructors and for platform developers. An important aspect of OLEs regards their socio-emotional elements, which often pose significant challenges. Social presence makes learners feel more engaged, while their experience tends to be more satisfying. According to Weidlich and Bastiaens (2019), Moodle plugins such as "Meet the Students", "Course Contacts", and "Dialog" establish a significantly more sociable learning environment; however, no effects of these plugins have been observed in terms of learning achievements. In addition, feedback is both important and useful for learners. For example, Yilmaz and Yilmaz (2020) analyzed the impact of different forms of feedback (i.e., text, image, video) on the transactional distance perceptions and critical thinking skills in online discussions. Their findings showed that the form of feedback had significant effects on transactional distance perceptions, wherein the lowest transactional distance perception was found for video feedback, followed by image and text feedback. By contrast, critical thinking skills were not influenced by the form of feedback.
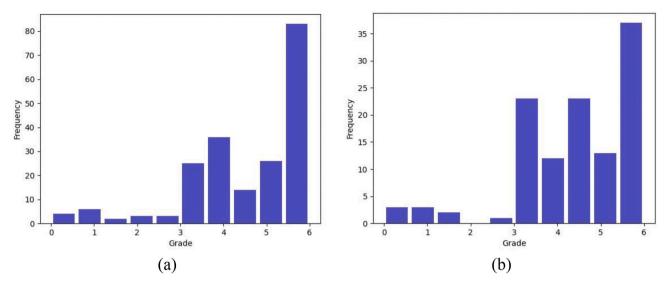
Fig. 1. Histograms of student grades in a) 2018–2019 academic year and b) 2019–2020 academic year.

Different online learning platforms appeared throughout the years to support both students and tutors in their learning activities. Depending on the course, its purpose, and how students or tutors want to attend or teach, various types of learning platforms have attempted to cover a wide range of requirements. Moodle (Dougiamas & Taylor, 2003) is one of the most popular OLEs. Its theoretical perspective of social constructionism and connected knowing supports both students and teachers in multiple learning activities such as publishing course content; posting and viewing information (e.g., deadlines, events); submitting and viewing grades and homework; and supporting discussions, collaboration, group work, and private communications amongst students. Moodle is free, easy to use, multilingual, robust, secure, scalable to large audiences, highly flexible, and fully customizable (Moodle Docs, n.d.).

2.2. Predicting course success

One line of OLE research has focused on predicting course completion and success based on various data that can be automatically accessed while the students interact within the environment. Clickstream data provide details about students' interactions with the course content, discussions forums, and assignments. Common measures include the various types of actions students can take, the number of different actions they perform, how often a certain action is done, the time when certain actions are performed (relative to others), discussion forum interactions, and assignments attempted. Multiple studies have demonstrated that clickstream data representing students' behaviors and involvement during online courses are predictive of outcomes such as grades and completion rates (Li, Baker, & Warschauer, 2020; Seaton, Bergner, Chuang, Mitros, & Pritchard, 2014; Sharma, Jermann, & Dillenbourg, 2015). In general, these studies have demonstrated that clickstream measures were significantly associated with students' self-reported measures (e.g., time management), with reported algorithmic prediction accuracies of course performance surpassing baseline measures such as students' self-reports.

Another approach is to consider the language that students generate in discussion boards. The students' responses and statements on various conversation threads can be analyzed using various types of Natural Language Processing (NLP) tools (McNamara, Allen, Crossley, Dascalu, & Perret, 2017). A common approach is to analyze students' opinions about a course and its activities using sentiment analysis tools. These are tools that focus on the emotional valence of the words used by the students. For example, Wen, Yang, and Rose (2014) examined the statements that students expressed in their daily posts. They reported that

students who used words related to motivation and personal pronouns had a lower risk of dropping out of the course.

Crossley et al. (2015) were the first to go beyond the use of click-stream and sentiment analyses in predicting online course success. They examined the language in students' posts on the course discussion board, including features related to lexical and syntactic properties, text cohesion, syntactic similarity, as well as features related to sentiment. Their results indicated that those who completed the course tended to be better writers, used a wider variety of words, and wrote longer and more cohesive messages. The importance of language knowledge and its use to course completition was an important finding, particularly because this was a course on educational data mining and none of the assignments involved writing.

Crossley et al. (2016) combined click-stream and NLP features of the students' discussion board posts. They found that including the click-stream variables provided an improvement over models based solely on linguistic features of about 10%; accurately predicting student completion rates with an accuracy of 78%. Successful students submitted their assignments earlier and interacted with the course more often (e.g., viewed the lectures). Crossley, Karumbaiah, Ocumpaugh, Labrum, and Baker (2019) conducted a study examining performance in a math course. They combined students' demographic information with the click-stream and linguistic variables from NLP tools designed to measure lexical sophistication, text cohesion, and sentiment, to predict success rates in solving math problems in an online tutoring system. Students having a more diverse vocabulary, a more cohesive and sophisticated language, with a more intense online activity (i.e., increased total time spent and number of entries into various modes logged by the system) were more successful at advanced level math problems.

3. ReaderBench framework

3.1. ReaderBench and Cohesion Network Analysis (CNA)

This study builds on prior research by combining various sources of information about students' performance, including linguistic and semantic features of their posts, clickstream data, and aspects of their interactions. *ReaderBench* is an open-source framework (http://readerbench.com/) that computes a wide variety of textual complexity indices, including lexical, syntactic, as well as semantic and discourse centered facets of dialog (M. Dascalu, Crossley, McNamara, Dessus, & Trausan-Matu, 2018). Cohesion Network Analysis combined with textual complexity indices in *ReaderBench* were used to predict student course grades in a Romanian Moodle course with a Mean Average Error
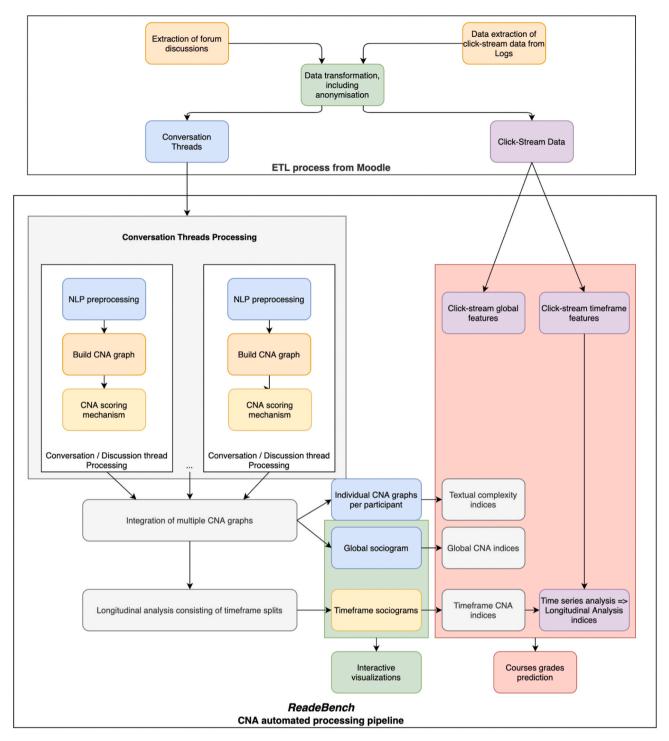
Fig. 2. *ReaderBench*: Processing pipeline for predicting grades and generating interactive visualizations.

of 0.4438 on a 6-point scale (M.-D. Dascalu et al., 2020). This study also introduces a Recurrent Neural Network architecture to perform time series analyses and a consolidated prediction of student performance, coupled with updated visualizations to better highlight different behavioral patterns. These visualizations are designed to support instructors in their understanding of students' performance in these online environments.

Cohesion Network Analysis (CNA; M. Dascalu et al., 2015) combines advanced NLP approaches with Social Network Analysis (SNA; Scott, 2017) to analyze and provide an in-depth view of discourse structure centered on text cohesion. SNA represents and examines social

structures using graph theories; CNA is closely correlated to SNA as it provides equivalent indices to evaluate participation using network graphs that use estimates of discourse cohesion to simulate information exchanged between participants. CNA improves SNA because it considers semantic cohesion based on Natural Language Processing (NLP) when modeling students' interactions. As a consequence, CNA considers both students' interactions and discourse content to conduct an in-depth analysis of students' interactions.

Within the CNA, cohesion is computed using various similarity measures from different semantic models, namely: Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), Latent Dirichlet Allocation

Table 1

CNA participation and initiation indices.

CNA index	Description
CNA Participation indices	
contribution score	Sum of contributions' scores for each participant
indegree	Sum of in-edges from the cohesion graph predictive of collaboration with other members
outdegree	Sum of out-edges from the cohesion graph indicative of active participation
social knowledge building (KB)	Sum of the edges between a given participant and other participants from the community, reflecting collaboration
closeness centrality	Sum of the lengths of the shortest paths between a node (i.e., participant) and all other nodes in the graph (the more central a node is, the closer it is to all other nodes)
betweenness centrality	Shortest paths that pass through a node in the participant graph
eigenvector centrality	Reflects the participant's influence in the community using eigenvalues computed on the participant graph
CNA Initiation indices	
new threads	Conversation threads initiated by a given participant (i.e., the selected participant had the first post)
overall score	Sum of the contribution scores of all utterances from initiated discussion threads by a given participant
average length	Average count of contributions per initiated discussion thread

Table 2

Longitudinal analysis indices.

Longitudinal analysis index	Description
Features applied on specific timeframe CNA indices	
M & SD	Mean (M) and standard deviation (SD) of the considered timeframe CNA index, within all timeframes
local extreme points	Count of timeframes for which the inflection of the CNA index changes (e.g., an increase of participation followed by a decrease); this index reflects the degree of monotony within the evolution of each participant or whether spikes are encountered
slope	Degree of the slope, indicative of students' involvement, obtained after applying a linear regression to the time series; a slope greater than 0 means that the students were more actively engaged; a slope of zero denotes a uniform involvement, whereas a negative slope indicates that students have lost their interest and have a lower participation in subsequent timeframes
Global longitudinal analysis indices	
activity	Percentage of timeframes in which the given participant had a least a contribution
M & SD recurrence	Recurrence is computed as the distance between timeframes when the student had at least one contribution; recurrence increases if breaks in online engagement are encountered (e.g., recurrence is 1 if a student posts every two weeks, whereas a value of 0 means perfect regularity and weekly posts)

(LDA; Blei, Ng, & Jordan, 2003) or word2vec (Mikolov, Chen, Corrado, & Dean, 2013). The cohesion graph represents a multi-layered structure, consisting of different nodes and the links between them, and it can be used as a proxy for the semantic content of the discourse (M. Dascalu, 2014). The cohesion graph consists of a central node, which represents the conversation's thread. Then, the central node is divided into contributions, which are further divided into sentences and words. Links are built to compute a cohesion score that denotes the relevance of a contribution in a conversation, or the impact of a word in a sentence or contribution. The graph also includes explicit links added by the participants in the conversations, such as "reply-to". Besides predicting collaboration (M. Dascalu, McNamara, et al., 2018) and course grades (M. Dascalu, McNamara, et al., 2018), CNA was also successfully employed to predict blogger community response to newcomer inquiries via automated dialog assessment (Nistor, Dascalu, Serafin, & Trausan-Matu, 2018; Nistor, Dascalu, Tarnai, & Trausan-Matu, 2020).

The theoretical backbone of *ReaderBench* and CNA is *dialogism*, which emphasizes the importance of dialog and its apex, polyphony (Trausan-Matu, 2020). Accordingly, semantic overlap (i.e., cohesion) and the flow of information (i.e., polyphony) are essential for knowledge construction in Computer-Supported Collaborative Learning (CSCL; M. Dascalu et al., 2015). This paradigm characterizes the learning situation targeted in this study, where students collaborate both in formal activities on OLEs, but also in parallel, on other online environments. This second dimension occurs outside the teachers' control, for conversations about the lectures, homework assignments, and exams, in the absence of face-to-face discussions at faculty. The complex analyses performed by *ReaderBench* are based on the artifacts created during the learning sessions (Trausan-Matu & Slotta, in press), while Moodle provides access to a varied series of such artifacts. Cohesion Network Analysis is using a formal abstraction of the learners' dialogs, which are text artifacts generated during CSCL sessions, recorded by Moodle and used to provide complex statistical data and visualizations, following from the dialogism theory.

We have developed a *ReaderBench* plug-in to Moodle, which can be used to evaluate the online activity of students, illustrating the underlying nature of their behaviors and interactions. We examine the extent to which features generated by *ReaderBench*, including CNA applied on forum discussions, textual complexity indices, longitudinal analysis, and features derived from click-stream data, successfully predict student performance in an undergraduate course on Algorithm Design. More pointedly, we are able to examine differences in students' behaviors and interactions before and during COVID-19 in two consecutive yearly installments of an undergraduate course on Algorithm Design. To this end, we conduct an in-depth analysis of students' participation in Moodle in tight relation to the global pandemic, while identifying interaction patterns and predicting student grades based on their activities in online forum discussions and on click-stream data. Multiple CNA visualizations that consider text cohesion among students' posts are generated to observe the nature of various interactions patterns. Weekly snapshots help tutors better understand students' behavior, while associating their activities with course events such as deadlines, assignments, tests, and exams.

4. Method

4.1. Corpus

Our data was collected from two different Moodle instances from the 2018–2019 (*normal* conditions) and 2019–2020 academic years (*COVID-19* pandemic conditions) on a Moodle course from the second semester, held in Romanian and centered on Algorithm Design. The data included forum posts of students, lecturers and teaching assistants, and their online activities extracted from click-stream log data. The information collected from forum posts consisted of usernames, contributions' timestamps, reply-to links, and the actual texts from the posts.

In the academic year 2018–2019, a total of 202 students were enrolled in the course and wrote posts on the Moodle platform; students with no posts were disregarded as there were no textual traces to analyze. Students were guided by 3 lecturers and 16 teaching assistants, generating 118 discussion threads and 632 contributions. The partial grades that account for points awarded for activities throughout the semester (e.g., assignments, course project, answers during labs) were taken into consideration for this experiment. The students achieved partial grades that ranged from 0 to 8.65 (M = 4.82, SD = 1.66) out of 6 (the maximum points for the activity during the semester, the other 4 points from a maximum 10 being allocated for the final exam). Bonus points were awarded (thus arguing for values greater than 6), but all grades were capped to 6 in these analyses (M = 4.42, SD = 1.41; see Fig. 1a). The course lasted for 14 weeks (i.e., between February 18, 2019 and May 24, 2019), followed by 3 weeks of exam sessions. All discussions were in the Romanian language.
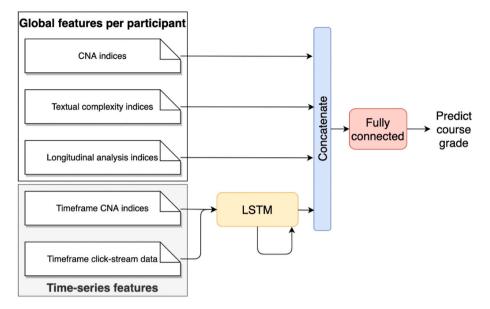
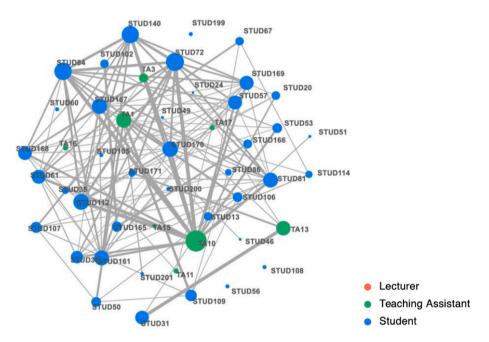**Fig. 3.** RNN architecture for predicting course grades.

Fig. 4. Force-directed graph for the 13th week 2018–2019 school year.

Within the next academic year (2019–2020), 117 students wrote posts on the Moodle platform, divided as in the previous year into three student-cohorts, guided by the same 3 lecturers and 15 teaching assistants. In total, 135 discussion threads and 535 contributions were generated. The partial grades with points gathered throughout the semester were taken into consideration for this experiment. Students achieved partial grades that ranged from 0 to 7.61 (M = 4.53, SD = 1.55) out of 6. Additional bonus points were also awarded and the final grades considered caped values (M = 4.62, SD = 1.42; see Fig. 1b). This normalization was performed because multiple bonus points (e.g., participation to contests, such as ACM International Collegiate Programming Contest) were awarded on different criteria in the two academic years, and a comparative scoring baseline was required to build transferable models across the two years. The mean and standard deviation values are comparable between the two years. The course lasted for 14 weeks (i.e., between February 17, 2020 and May 22, 2020) and it

was also held in Romanian.

4.2. ReaderBench: processing pipeline for forum discussions and click-stream data in moodle

Our approach is grounded in Cohesion Network Analysis combined with Machine Learning techniques, and it is used to evaluate and model students' participation and interactions. In contrast to the previous studies performed by M.-D. Dascalu et al. (2020), M. Dascalu, McNamara, et al. (2018) and Crossley et al. (2016), we introduce an integrated pipeline that accounts for all types of indices (CNA, time series and textual complexity), actions derived from clickstream logs, and a neural network for predicting course grades. We consider students' behaviors as revealed by actions within the OLE (e.g., viewing assignments, completing assignments, posting comments), as well as social interactions and the semantic content of their online contributions. Fig. 2
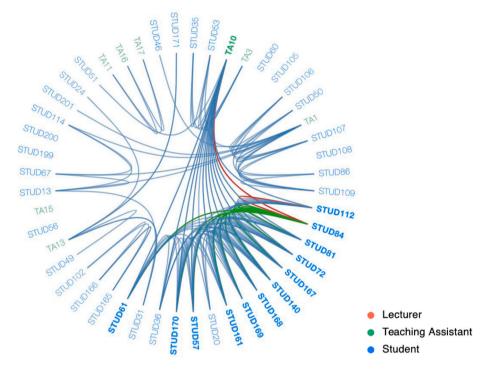
Fig. 5. Hierarchical edge bundling view for the 13th week 2018–2019 school year.
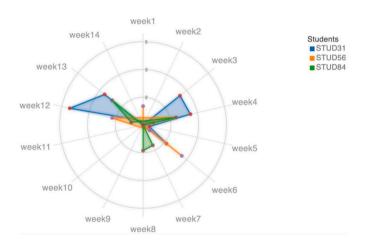

Fig. 6. Radar chart view illustrating contribution scores for the 3 most active students during the 2018–2019 school year.

presents the automated processing pipeline which includes two important stages. The first stage consists of an ETL (Extract Transform Load) process that starts with the collection of click-stream data and forum discussions from the Moodle platform which were exported from the relational database (i.e., MariaDB). Discussion threads from forums were extracted with corresponding information about each post (i.e., username, date, post, reply).

Click-stream logs were also extracted from the database using queries on specific tables and the logs included information for each user in terms of authentication actions (e.g., login) or submission of specific assignments. Next, data was anonymized, and two datasets were created: *conversation threads* (including the hierarchical structure) and *click-stream data*.

The second stage is centered on the automated processing pipeline from the *ReaderBench* framework, which has as input data the two datasets generated in the first stage. The two datasets, *conversation threads* and *click-stream data*, follow separate processing flows, detailed

below.

Each conversation thread undergoes processing comprised of three steps. First, an NLP pre-processing pipeline is applied, which includes diacritics restoration (Masala, Ruseti, & Dascalu, 2020), followed by tokenization, part of speech tagging, dependency parsing, stop word elimination, and lemmatization, all using a spaCy model trained for Romanian.[1] Even though deep learning methods for NLP usually do not require any preprocessing of the input text, these architectures are not applicable when processing these conversations. The structured nature of the forum conversations and the small size of the corpus are more suitable for smaller neural models that use handcrafted features. In addition, specific CSCL heuristics were applied, such as: merging adjacent contributions per participant (if a participant had several adjacent utterances within 30 s, merge the contributions into a larger one) and identifying reply-to links. Second, a cohesion graph is built to serve as a proxy for the underlying semantic content of each discussion thread. Third, a scoring mechanism is applied to determine the importance of each contribution within each discussion thread, which consists of a modified Page Rank algorithm (Brin & Page, 1998) applied on the CNA graph (Cioaca, Dascalu, & McNamara, 2020). Our approach is inspired from the TextRank algorithm initially developed for extractive summarization (Mihalcea & Tarau, 2004) – a task also suitable for online conversations in which the aim is to identify the most central contributions.

Afterwards, the CNA graphs of each discussion thread from all forums are integrated to create a graph corresponding to the entire community. Further, individual CNA graphs per participant (i.e., collections of all contributions pertaining to a given participants that are also represented using CNA graphs) and the global sociogram (network graphs depicting the interactions between participants) are computed. The individual CNA graphs per participant are further used to compute the textual complexity indices for each participant, indicative of their writing style.

---

[1] https://spacy.io/models/ro#ro_core_news_lg, last accessed on 5th January 2021.
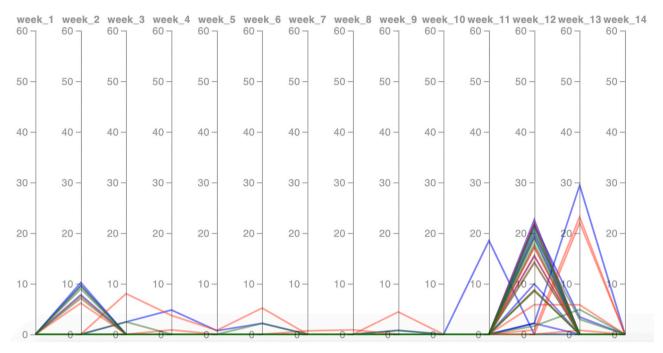
Fig. 7. Parallel view for the 10 most active students from the 2018–2019 school year illustrating trends in the community.

Global CNA indices are computed based on the global sociogram as SNA measures applied on the global participant graph (see Table 1). The CNA indices are divided in two categories: *participation* (M. Dascalu, McNamara, et al., 2018) and *initiation* indices (Nistor, Dascalu, & Trausan-Matu, 2016). Initiation CNA indices refer to the community's activity that occurs as a result of initiating a new discussions thread. Unlike previous studies, we combine both participation and initiation indices.

In addition, a longitudinal analysis (LA) is performed based on the integrated CNA graphs of each conversation thread (Sirbu et al., 2019). Timeframe sociograms are generated as splits of the entire community for each week (i.e., filtering only contributions that occurred within the imposed timeframe) and are used to observe how the community evolves from one week to the following one (Crossley, Dascalu, Baker, McNamara, & Trausan-Matu, 2017). Further, timeframe CNA indices are computed to analyze the participation and initiation of each participant within each course week. Subsequently, a time series analysis is applied on the timeframe CNA indices and additional longitudinal analysis indices are generated (Allen et al., 2016). Table 2 introduces the features from time series analysis (Crossley et al., 2017) considered in our prediction model which are split into two sub-categories: a) LA features applied on specific timeframe CNA indices and b) global LA features that are independent of any CNA index.

4.3. ReaderBench: predicting course grades

All indices including CNA, textual complexity, longitudinal analysis, as well as time series applied on timeframes and click-stream data are used to predict students' course grades using various machine learning algorithms. All previous indices provide valuable insights in terms of participation and collaboration (CNA indices), stylometry (textual complexity indices), online activity (click-stream data), as well as regularity and evolution in time (longitudinal analysis).

Our model consists of a Recurrent Neural Network (RNN) with LSTM cells (Hochreiter & Schmidhuber, 1997) that combines global features (CNA, textual complexity, and LA indices) with a time series analysis on timeframes in order to predict student grades. This architecture (see Fig. 3) was selected because it provides state of the art results on time-series analyses used to provide various forecasts based on historical data (Gers, Eck, & Schmidhuber, 2002; Siami-Namini, Tavakoli, & Namin, 2018). In our case, the inputs are weekly student activities derived both from CNA indicative of online participation and click-stream data reflecting overall interactions with Moodle.

The LSTM network receives a window of fixed size (14 consecutive time intervals, one corresponding to each week of the course) as input, each element being represented by the timestamp features for a given step, both timeframe CNA indices and timeframe click-stream features. The output from the last LSTM cell contains a representation for the entire course computed using weekly indices. This representation is subsequently concatenated with global features for each student (i.e., CNA, textual complexity, and longitudinal analysis indices) and the neural network computes the course grades using a fully connected layer. For each training example, the RNN minimizes the squared distance between its prediction at the final step and the correct student grade.

4.4. ReaderBench: generating interactive visualizations

Interactive visualizations are generated using the global and timeframe sociograms, to highlight the interaction between participants and to depict the evolution of the community from one week to the next. Thus, multiple types of visualization are rendered to depict students' evolution, behaviors, and interaction patterns using the d3.js library (https://d3js.org/). The web application was built using Angular 6 framework, while various JavaScript libraries were integrated to create the interactive sociograms.

The first view consists of a *force-directed graph* (https://observablehq.com/@d3/force-directed-graph) that illustrates a network graph in
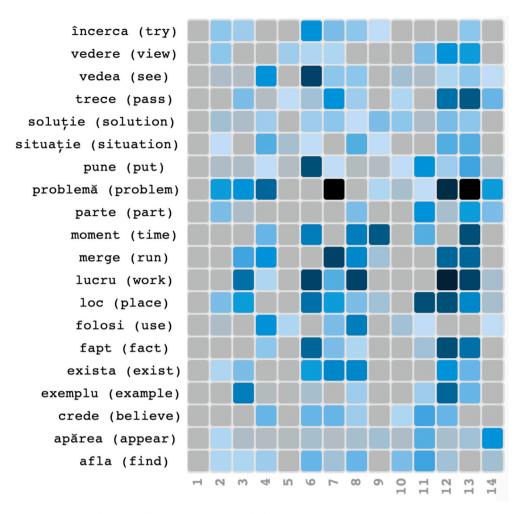
Fig. 8. Concept heat map illustrating most frequently discussed keywords during the 2018–2019 school year.

which the nodes represent the participants (i.e., students, teaching assistants, or lecturers), and the edges reflect the exchanged messages between participants. The size of the node is directly proportional with the sum of CNA *indegree* and *outdegree* indices. The width of the edges is proportional to the cumulative semantic links of inter-exchanged messages between the two participants representing the nodes. The nodes are colored depending on their role: orange – *lecturer*, green – *teaching assistant*, blue – *student*. For example, Fig. 4 presents the community during the 13th week of classes in the school year 2018–2019 – one week before the end of the course. This visualization illustrates which teaching assistants are more actively communicating with the students (i.e., TA1, TA10, and TA13), and which students are more actively communicating (e.g., STUD72) or less active students (e.g., STUD51). The links, however, are not just frequency of communication – they represent the semantic overlap within and across messages. We use these visualizations to compare communities as a function of school year and across time when we discuss the research findings for this study.

Second, the *hierarchical edge bundling* view (https://observablehq. com/@d3/hierarchical-edge-bundling) depicts connections between participants using a radial view (see Fig. 5, generated for the same timeframe as Fig. 4). This type of view bundles adjacent edges to decrease the clutter usually present in complex networks. The dependencies between participants are displayed in a radial manner, while the participants are colored based on their corresponding role. On a mouseover event on a participant's name, the *incoming* (green color) and *outgoing* (red color) links are highlighted, while the participants' name is bolded. The incoming links or reply-to messages denote collaboration, while the outgoing links are indicative of active participation. Weekly

sociograms were also generated using these first two visualizations to observe the evolution of participants from one week to another, as well as the impact of different deadlines to various activities on the involvement of participants in the forum discussions. Thus, weekly snapshots based on CNA can support tutors to better understand students' behaviors in association with course events such as deadlines, assignments, tests, and exams.

Third, the *radar chart* (https://www.d3-graph-gallery.com/spider. html) is a two-dimensional view designed to plot one or more series of values over multiple variables (weeks in our case). Each variable has its own axis, and all axes are joined in the center of the chart. Weekly evolutions of contribution scores for the first 3 students ranked by overall contribution scores were rendered in a radar chart view to observe the variation between students' degree of participation and to compare their activity. Using this analysis, tutors could take actions to stimulate students to become more engaged and participate more in the discussions. Fig. 6 presents the evolution of contribution scores across all timeframes for the most active students from the first academic year (i. e., 3 most active students having the highest overall contribution scores) on which a logarithm scale was applied. Students posted more in weeks with deadlines (4, 6, 12, and 13) in contrast to other weeks.

Fourth, a *parallel coordinates chart* (https://www.d3-graph-gallery. com/parallel) shows the evolution of CNA *indegree*, *outdegree* and *social knowledge building* indices per course week. These CNA indices are considered to concurrently highlight student behaviors in terms of cohesive links reflective of active participation (outdegree), reactions from other students (indegree), and collaboration effect generated within the discussion (social KB). Blue lines represent the evolution of

Table 3

CNA, textual complexity, and longitudinal analysis indices that exhibited significant differences as a function of academic year.

Feature	M 2018–19	SD 2018–19	M 2019–20	SD 2019–20	U	p
TC: M word polysemy count (# senses)	2.454	1.725	3.883	1.226	5966	<.001
TC: M unique verbs per contribution	3.556	3.699	7.622	5.289	5977	<.001
TC: M syllables per sentence	18.275	16.649	29.277	15.692	6279	<.001
TC: M verbs per sentence	1.748	1.812	2.962	1.716	6416	<.001
TC: Word entropy	1.898	0.572	2.298	0.362	6575	<.001
TC: M sentence parse tree depth	3.794	1.041	4.542	1.156	6836	<.001
TC: M adverbs per contribution	2.129	2.851	5.267	5.890	7257	<.001
TC: M punctuation signs per contribution	3.226	3.786	5.910	4.918	7334	<.001
CNA: Eigenvector centrality	0.030	0.059	0.045	0.064	7881	<.001
LA: SD contribution score	0.491	0.471	0.729	0.544	7891	<.001
LA: M contribution score	0.181	0.243	0.281	0.319	8068	<.001
TC: M word path hypernym tree	7.640	5.715	10.328	4.008	8307	<.001
TC: M word syllable	1.536	0.366	1.723	0.206	8572	<.001
CNA: New threads overall score	5.126	6.051	8.863	11.952	8862	<.001
TC: SD word length per sentence	2.647	1.436	3.098	1.107	9298	.001
TC: M word difference between flectional form and lemma	0.351	0.107	0.412	0.224	9420	.001
TC: M punctuation signs per sentence	1.228	0.771	1.314	0.693	9779	.005
LA: slope of contribution score	0.011	0.042	−0.006	0.054	9772	.005
TC: M adverbs per sentence	0.813	0.656	1.128	1.064	10036	.012
TC: M word average depth in hypernym tree	2.378	1.741	3.008	0.993	10057	.013
LA: SD recurrence	3.388	1.530	3.135	1.581	10595	.061
CNA: new threads average length	1.646	1.080	1.872	1.450	10747	.062
LA: slope of eigenvector centrality	0.000	0.003	0.000	0.004	10792	.098
LA: activity	0.108	0.073	0.118	0.084	11059	.124
LA: M eigenvector centrality	0.007	0.015	0.011	0.017	10912	.127
LA: local extreme points of eigenvector centrality	1.475	1.077	1.402	0.943	11449	.274

* CNA = Cohesion Network Analysis; TC = textual complexity; LA = Longitudinal Analysis; the features are presented as the function applied over a specific CNA index across all timeframes.

Table 4

Prediction results for the 2018–2019 academic year.

CNA	TC	LA	CS time series	LA time series	LSTM cell	Hidden	CV RMSE	Test RMSE	Test R^2
✓	✓	✓	✓	✓	16	16	.213	.229	.176
✓	✓	✓	✓	✓	8	16	.219	.221	.235
✓	✓	✓	✓	✓	24	16	**.218**	.222	.230
✓	✓	✓	✓	✓	24	8	.219	.229	.181
✓	✓	✓	✓	✓	24	24	.219	.218	.252
✓	✓	✓	–	–	–	16	.219	.227	.190
✓	✓	✓	–	✓	24	16	.221	.226	.200
✓	✓	✓	✓	–	24	16	.219	**.216**	**.265**
–	✓	✓	✓	✓	24	16	.223	.229	.179
✓	–	✓	✓	✓	24	16	.228	.249	.025
✓	✓	–	✓	✓	24	16	.219	.231	.162

* CNA = Cohesion Network Analysis; TC = textual complexity indices; LA = longitudinal analysis indices; CS = click stream; LA time series = RNN across timeframes; CV = cross-validation.

Table 5

Prediction results for the 2019–2020 academic year.

CNA	TC	LA	CS time series	LA time series	LSTM cell	Hidden	CV RMSE	Test RMSE	Test R^2
✓	✓	✓	✓	✓	16	16	.217	.209	.316
✓	✓	✓	✓	✓	16	24	.215	.206	.333
✓	✓	✓	✓	✓	16	8	.221	.225	.203
✓	✓	✓	✓	✓	8	24	.216	.209	.315
✓	✓	✓	✓	✓	24	24	.214	.211	.304
✓	✓	✓	–	–	–	24	.219	.208	.319
✓	✓	✓	–	✓	24	24	**.213**	.209	.316
✓	✓	✓	✓	–	24	24	.213	**.205**	**.341**
–	✓	✓	✓	✓	24	24	.222	.222	.225
✓	–	✓	✓	✓	24	24	.222	.240	.097
✓	✓	–	✓	✓	24	24	.213	.211	.302

* CNA = Cohesion Network Analysis; TC = textual complexity indices; LA = longitudinal analysis indices; CS = click stream; LA time series = RNN across timeframes; CV = cross-validation.
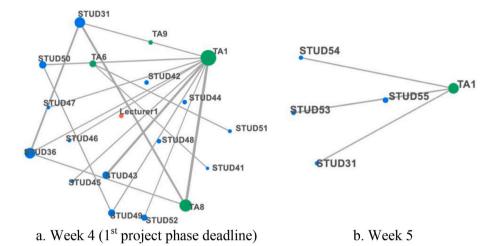
a. Week 4 (1st project phase deadline)   b. Week 5

Fig. 9. Force-directed graphs for the 4th and 5th weeks of the 2018–2019 school year.

a. Week 7 (1st homework deadline)   b. Week 8 (2nd project phase deadline)   c. Week 9
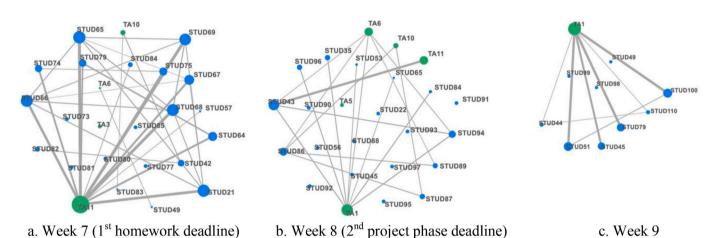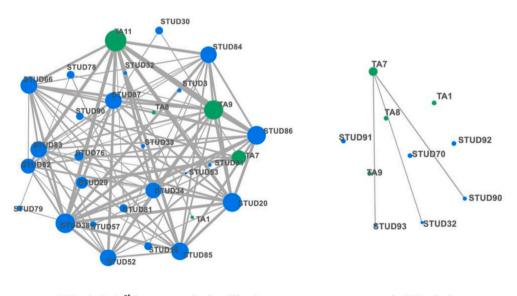
Fig. 10. Force-directed graphs for the 7th, 8th and 9th weeks of the 2018–2019 school year.

a. Week 8 (1st homework deadline)   b. Week 9

Fig. 11. Force-directed graphs for the 8th and 9th weeks of the 2019–2020 school year.
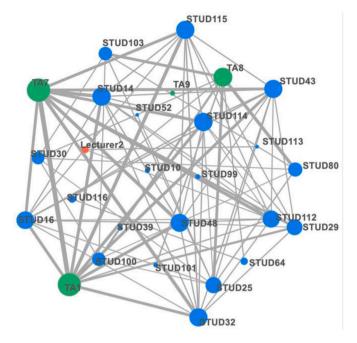
Fig. 12. Force-directed graph for the 13th week 2019–2020 school year.

social knowledge building indices, red lines the evolution of outdegree indices, while green lines represent the evolution of indegree indices. The evolution of students' participation and collaboration can be visualized using this parallel view (see Fig. 7). For example, the second week exhibits a high collaboration among students (the social KB scores for top overall 10 students are all colored in blue). Week 12 has high values for all 3 considered dimensions for the most highly engaged students, while week 13 is dominated by a single student (only 1 participant who has a high social KB - blue and a high outdegree – red, while responding to several peers). Behavioral trends can be observed during the course: students would rather participate in discussions than collaborate, changes take place in the weeks when students have tests or homework deadlines (i.e., weeks 4 and 13).

In addition, we examined which keywords are most frequent in the discussions from Moodle and we represented the most frequently discussed concepts using a concept heatmap. Words represent the rows within this grid view, columns represent consecutive weeks, whereas colors code the frequency of each word (i.e., grey means no occurrence, light blue denotes a low usage, while dark blue reflects a high frequency). The 20 most frequently discussed keywords globally were extracted, which include only nouns and verbs, excluding Romanian modal verbs. A *heatmap* visualization (https://www.d3-graph-gallery.com/heatmap) was used to represent frequent keywords throughout all weeks (see Fig. 8). For example, keywords related to the exam and solutions to assignments (such as, "afla" - eng. "find", "soluție" - eng. "solution", "trece" - eng. "pass") were intensively used during the course in the academic year 2018–2019, while "lucru" (eng. "work"), "problemă" (eng. "problem"), "merge" (eng. "run"), "loc" (eng. "place") were used in the weeks with homework deadlines (i.e., weeks 7 and 13).

4.5. Analyses

The machine learning analysis to predict students' course grades included a training set ($n = 182$ for the 2018–2019 school year and $n = 97$ for the second academic year) and a test set comprised of 20 randomly chosen students (for each of the two academic years). A 10-fold cross-validation was performed on the training set to identify the hyper-parameters for the regression models using a grid search. Student grades were capped at 6 points and normalized on a [0, 1] scale (i.e., division by 6), denoting students' activity throughout the 14 weeks and

before the exam session. Textual complexity indices were filtered in order to ensure linguistic coverage (i.e., at least 20% of values for all participants had to be non-zero). All features were checked for multi-collinearity using a Pearson correlation above 0.9; the most predictive features were retained in follow-up analyses. The reported metrics are Mean Average Error (MAE; the distance in absolute value between the predicted score and student's normalized score) and R^2 denoting the variance explained by our model.

All CNA and longitudinal analysis indices were checked for normality using the Kolmogorov–Smirnov test. If the test was statistically significant, we attempted to normalize the feature using a logarithm scale because most CNA features exhibit a long-tail distribution when accounting for participation in online communities. If the logarithmic index was still not normally distributed, a non-parametric Mann-Whitney U test was employed to compare the two student samples from the consecutive instalments of the course; otherwise, one-way ANOVAs were used to evaluate the differences between the group means.

5. Findings

5.1. CNA, textual complexity, and longitudinal analysis: before and after COVID-19

We explore differences in students' behaviors and interactions before and during COVID-19 using the CNA, textual complexity, and longitudinal analysis indices that entered the analysis after normality and multi-collinearity checks. After the initial check for linguistic coverage and multi-collinearity, only 26 indices remained after the statistical pruning. All CNA indices exhibited positively skewed distributions and all indices rejected the null hypothesis from the Kolmogorov–Smirnov test. Thus, all remaining indices were evaluated using the Mann-Whitney U test (see Table 3).

As expected, a significant increase in online activity is observed (i.e., higher CNA contribution scores) during the 2019–2020 academic year impacted by COVID-19. In addition, considerably more threads were initiated, and the overall network is more connected (i.e., higher CNA eigenvector values).

Textual complexity indices denote a more elaborated and sophisticated discourse in the second academic year. This can be viewed at multiple levels: superficial (i.e., more punctuation signs at sentence and contribution levels), word level (i.e., longer words in terms of syllables, increased word polysemy counts, more paths in potential hypernym trees generated by more senses, longer word inflections, a more diverse vocabulary reflected in word entropy), morphology (more verbs and adverbs at sentence and contribution levels denoting more actions and corresponding descriptions) and syntax (i.e., more complex sentences in terms of the parse tree depth).

Longitudinal analysis indices also reveal interesting results. An increased activity towards the end of the semester can be observed in the 2018-19 academic year (i.e., the overall slope is positive). In contrast, the COVID-19 outbreak generated an off-balance, a drastic increase in participation, that afterwards decreased towards the end of the semester (i.e., the slope was close to zero or negative, with higher standard deviations). These trends are described and visualized in greater detail in the following sections.

5.2. Course grades

The prediction of students' course grades was performed using the RNN architecture described in the Method section with various configurations of considered indices. The hyperparameters of the prediction model were selected based on the root mean squared error (RMSE) using 10-fold cross-validation. For the performance on the test partition, the RMSE and the R^2 metric were reported. To account for the small size of the dataset and the high variability of the results, the process was repeated 10 times and the average was taken into consideration. Several
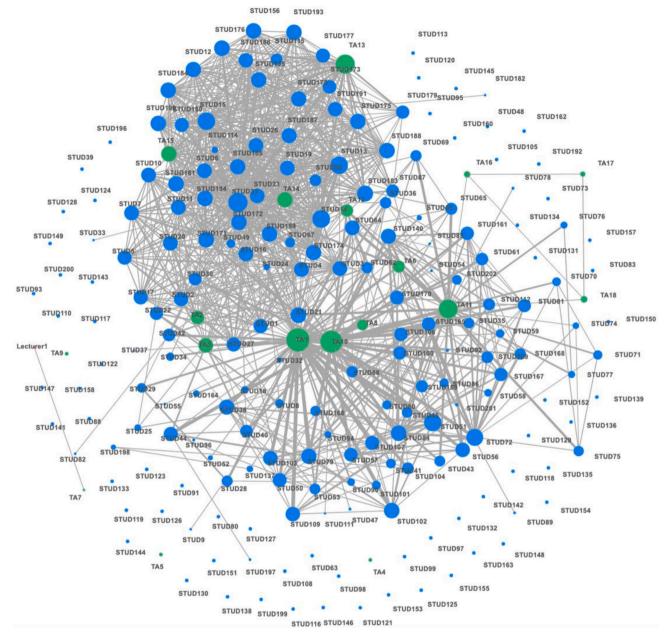
Fig. 13. Participants' CNA graph for the entire 2018–2019 school year.

experiments were performed varying the cell size of the LSTM units, the size of the final hidden layer, but also selecting different subsets of features.

Models were trained individually for each academic year. The models for both academic years were trained for a maximum of 2000 epochs with early stopping based on RMSE on the training loss, with a patience of 200 epochs. A dropout layer of 0.2 was introduced before the last hidden layer to reduce overfitting and improve generalization. These parameters were chosen based on the performance obtained on cross-validation.

There were five components of the model (CNA, textual complexity, longitudinal analysis - LA, click-stream time series, LA time series). To assess the importance of each component, we conducted an ablation study in which we tested various combinations of components and their contribution to explaining the variance in students' final scores in the course. An additional setup considers the removal of both click-stream and LA time series, denoting a configuration that uses only CNA, TC, and LA indices with no information from the RNN on weekly data.

The cross-validation (CV) and test results using various configurations for the first academic year are presented in Table 4, while the ones corresponding to the second year are included in Table 5. The first five entries introduce different configurations of the RNN hyperparameters (i.e., number of LSTM cells and the dimension of the last fully connected layer). The most predictive initial configuration in terms of CV RMSE was used in subsequent setups, where various configurations were experimented by removing categories of indices and observing the overall performance of the model.

For the 2018–2019 academic year, Table 4 demonstrates that the model that includes all components except for the CNA weekly features, and using an LSTM cell size of 24 and a hidden layer size of 16, explained the largest percent of variance ($R^2 = 0.27$). For the subsequent year, during COVID-19, the best performance was obtained with the same features, but with a hidden layer size of 24 instead of 16. This model obtained a higher R^2 of 0.34. Global CNA indices can be perceived as sums of weekly CNA indices from the longitudinal analysis; thus, it is not really surprising that they are not always useful in the RNN architecture.
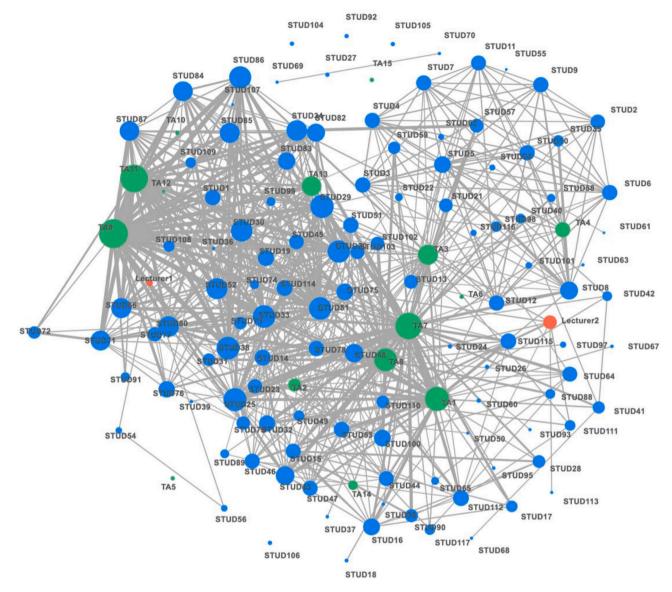
Fig. 14. Participants' CNA graph for the entire 2019–2020 school year.

A higher explained variance in the second academic year is justifiable, given the larger amount of collected data (i.e., more posts and a denser network) that helps create a more predictive model. In both academic years, the removal of textual complexity indices leads to the largest decrease in explained variance.
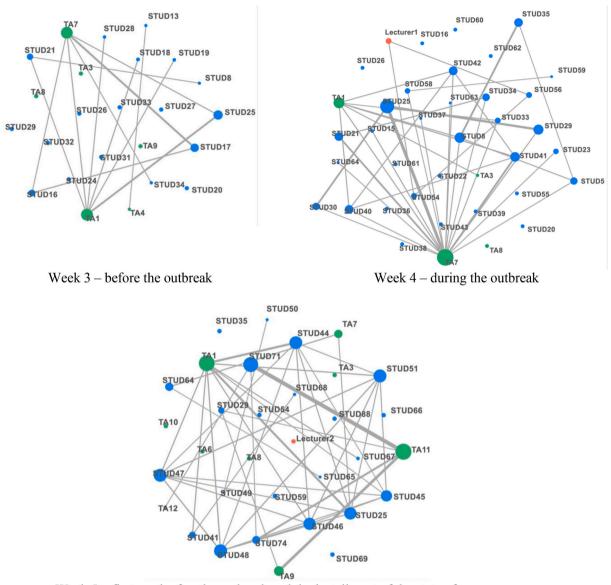
The best two configurations from the first year (normal conditions) were also tested on data from the second year in order to assess the model's generalizability. The configurations were trained on the whole dataset from the first academic year, and evaluated on the whole dataset from the second year; the two best models obtained an R^2 of around 0.13, thus arguing that the model trained on the first academic year can be still applicable for the second one (with a drastic decrease of performance from .34 to .13).

5.3. Qualitative description and analysis

In this section, we provide a qualitative analysis of moving to the purely online Algorithm Design course as a function of COVID-19 by considering impressions of the course lectures. The Moodle platform has been used for the Algorithm Design course for more than 12 years. Before the COVD-19 pandemic, it was used as a repository for course slides and other support documents, for laboratory documentation, and for announcements and discussions on a forum. The latter was seldom used by students, and even then, mainly for complaints, clarifications, or follow-up questions related to assignments. The discussion forum was used for debates on laboratory topics only in a few isolated occasions. The course lectures were in an amphitheater, combining slide projections and writing on a blackboard. The laboratories were in small groups of 12–15 students, each student having access to a computer.

After March 15, 2020, due to the COVID-19 pandemic, all learning activities went online. Lectures were given online on the Microsoft Teams platform, from home, using a laptop with an incorporated camera. The writing on the blackboard was replaced by additional slides and occasionally the professor (online) wrote with a stylus on them, or used Google Jamboard (https://jamboard.google.com) or OneNote (https://www.onenote.com) for additional notes, proofs, explaining solutions. The online lectures were recorded and uploaded on Moodle. After each lecture, professors were required to upload on Moodle a set of tests, which students had to complete by the end of the next day.

The observed effects of the transfer to online lectures that could potentially affect students' learning can be classified into three groups: changes in the lectures' presentation, changes in students' activities during the semester, and changes during the exam. Due to the online presentations, several non-verbal communication channels were no

Week 3 – before the outbreak

Week 4 – during the outbreak

Week 5 – first week after the outbreak and the installment of the state of emergency.

Fig. 15. Forum participation before and after outbreak caused by the COVID Pandemic.

longer available, specifically professor movement in front of students, gestures, and the alternation of slides presentation and writing on blackboard. The latter is useful for a course on algorithm design because it involves writing algorithms, investigating alternative solutions (involving erasing parts of the written text on blackboard and re-writing), which, in our opinion is better for understanding some subtleties than using only slides. A beneficial feature of the online lectures (with a laptop and incorporated camera) may be that the facial expressions of the professor are much more visible. The fact that students may see from home the lecture may be an advantage but also a disadvantage (potentially being distracted by other activities).

A few positive aspects of the online lectures were that students had the recordings and could access them later, during the semester or before exams. Another positive aspect may have emerged from the mandatory online tests after each lecture, which forced students to be attentive during presentations and subsequently reflect on their content. However, it can be assumed that some students simply completed the tests without a great deal of interest, and only for the purpose of being marked as present.

The exam was also online, on Moodle, consisting of the same three

parts as the previous years, when the exam was written on paper in amphitheater, without any sources of documentation. Students participated in the online exam from home and, because it was difficult to control their activities, they could use any documentation, on a physical or online form (internet). The three parts consisted of the exam included a series of theoretical quizzes, a problem for which they wrote an algorithm on paper, took a photo and uploaded it onto Moodle, and completed several problems wherein they wrote the steps of the execution of graph algorithms studied during the course, which were also written on paper and uploaded after being photographed. Notably, some students complained there were some difficulties when uploading the photos.

5.4. Visualizations: before and after COVID-19

In this section, we describe and analyze the sociograms generated by *ReaderBench* based on the CNA and LA indices. The visualizations were selected to exhibit different patterns of interaction correlated to course events, as well as the COVID-19 pandemic. In both years, students received a project consisting of three phases and two homework
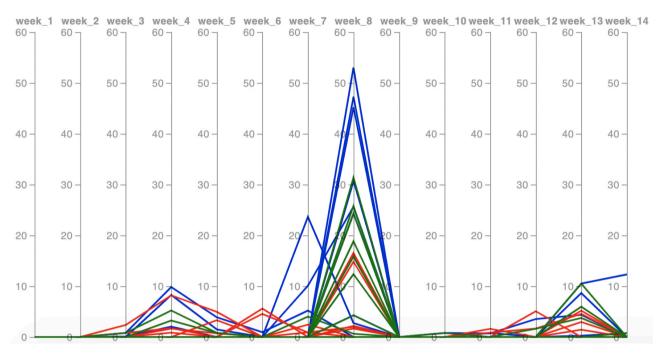
Fig. 16. Parallel view for 10 most active students from the 2019–2020 school year illustrating trends in the community.

assignments. Fig. 9 presents students' participation in weeks 4 and 5 from the academic year 2018–2019; students had the deadline for the first phase of the project in week 4, while their participation decreased drastically in the 5th week.

The same pattern was observed in weeks 7, 8, and 9 of the first academic year 2018–2019 (see Fig. 10); students had their first homework deadline in week 7. An intense collaboration took place with teaching assistant TA1 who was responsible for the first homework. In the next week (8), students had the deadline for the second phase of the project. More teaching assistants (TA1, TA6, TA11) were active and answered students' questions. Afterwards, a decrease in students' participation was noticed in the next week after the two deadlines.

Similarly, more students posted and collaborated in second academic year 2019–2020 during weeks in which they had deadlines, followed by a drastic decrease in the immediate follow-up week (see Fig. 11).

Compared to first academic year, students were more active in the second school year in weeks with deadlines, and collaborated more with each other, rather than solely with the teaching assistants. Strong connections were observed between students (see Fig. 10 b versus Fig. 11 a and Fig. 4 versus Fig. 12 for comparative views).

Overall, there were many students who were isolated on the forum in the first academic year, without interacting with peers or responding to other student inquiries. Also, only two teaching assistants were more active in all conversations (TA1, TA10), whereas lecturers were less engaged (see Fig. 13). In contrast, students were more engaged in the discussions from the 2019–2020 academic year (see Fig. 14), collaborated more, and there were fewer students who solely introduced isolated posts as new conversation threads. Overall, the network of participants is more connected in Fig. 13 as compared to Fig. 14, while more teaching assistants were actively involved in the conversations (TA1, TA7, TA9, and TA11), together with lecturers (Lecturer 2).

More conversation threads were created in the context of the COVID-19 pandemic and more students posted on Moodle, even in weeks when there were no deadlines. Romania entered a state of emergency on March 15, 2020 – the end of the week 4. Fig. 15 shows students' interactions in three weeks: before outbreak, during outbreak, and after outbreak. An intense activity was observed on Moodle in the weeks before and after the outbreak.

A longitudinal analysis was performed in order to evaluate involvement during the course and to examine differences between two academic years. The first five students with the greatest social knowledge building, indegree and outdegree indices were selected for both 2019 and 2020 years. In contrast to the evolution of social knowledge building, indegree, and outdegree CNA indices from the first academic year (see Fig. 7), students tended to post regularly, while collaboration was present more in the last weeks of the course. Students also posted regularly in the second academic year (see Fig. 16), with a spike in the 8th week when both a project phase and homework overlapped, and collaborated more with each other overall.

In addition to analyzing students' behavior and interactions, we examined which keywords were most frequently used in forum discussions. Some concepts' usage depended on the week (with or without homework deadlines), whereas others were frequently used throughout the entire academic year. Keywords like "problemă" (eng. "problem"), "exemplu" (eng. "example"), "situație" (eng. "situation"), "pune" (eng. "put") were intensively used during the course from the academic year 2019–2020, during the COVID-19 pandemic. In the weeks with homework deadlines (i.e., weeks 8 and 13), keywords like "lucru" (eng. "work"), "problemă" (eng. "problem"), "merge" (eng. "run") were also encountered (see Fig. 17). Compared with the previous academic year (see Fig. 8), keywords like "problem" and "work" were intensively used during the entire course, not only in the weeks with homework deadlines. Moreover, almost all 20 extracted keywords were intensively used in weeks 3, 4 and 5 (just before, and immediately after the outbreak).

6. Conclusions

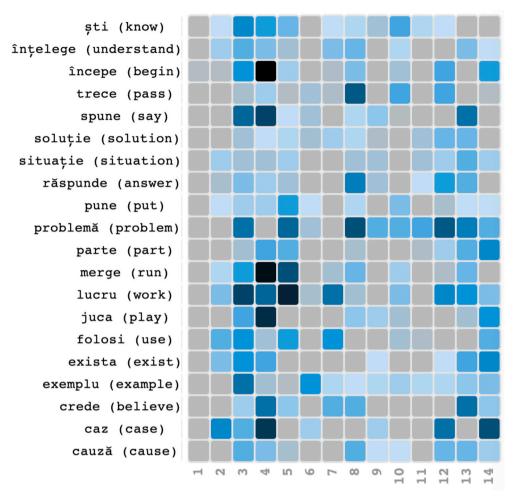Online learning environments are increasingly used by both students

Fig. 17. Concept Map 2020 - The most discussed topics.

and teachers, and their usage has increased during the COVID-19 pandemic. Within this study, we had three objectives: a) introduce the new version of the *ReaderBench* framework, grounded in Cohesion Network Analysis, which can be used to evaluate the online activity of students as a plug-in feature to Moodle, a popular open-source learning environment; b) evaluate the utility of *ReaderBench* by predicting students' course grades and illustrating students' interaction and behavior patterns using sociograms and c) compare students' behaviors and interactions before and during the COVID-19 pandemic. Data used in this study consisted of forum posts and click-stream data (logs) extracted from two consecutive yearly instances of an undergraduate course in Algorithm Design, conducted in Romanian using Moodle. A major benefit in the usage of the *ReaderBench* framework resides in its support for multiple languages, thus ensuring the method's applicability in multilingual settings.

In contrast to previous analyses (M.-D. Dascalu et al., 2020), the processing pipeline was extended with additional features and a RNN architecture, participation and initiation indices are combined, and new visualizations were introduced to provide more accurate predictions and to better highlight different behavioral patterns. In the current study, force-directed graphs and hierarchical edge bundling views were generated to illustrate the interaction patterns between students, teaching assistants, and lecturers. A longitudinal analysis was performed to evaluate involvement throughout the course and was integrated in the RNN model, while parallel views were introduced to observe the differences between the two academic years. Concept heatmaps depicting the first 20 most discussed keywords were extracted for each year, highlighting the COVID-19 outbreak that generated intense discussions,

but concurrently denoting resemblances in terms of homework deadlines.

The best prediction model for both academic years used all features accept the weekly CNA features and obtained R^2 scores of 0.27 for the first year and 0.34 for the second one. The model from the first academic year is partially generalizable for the second year, thus arguing for the specificity of student behaviors in the year of the outbreak. A significant increase in online activity was observed in the academic year 2019–2020 during the COVID-19 pandemic, with considerably more threads initiated, while the overall network is more connected. An increased activity towards the end of the semester was observed in the academic year 2018–2019 with lower fluctuations in participation across the semester. In contrast, the COVID-19 outbreak generated an off-balance, a drastic increase in participation, that afterwards decreased towards the end of the semester.

The transfer from face-to-face lectures, in an amphitheater, to full online lectures and laboratories generated two types of changes revealed by *ReaderBench*. The first category of changes was expected: the number of online activities and the length of discussion threads increased, the number of isolated students (that did not connect with others online) decreased. These are normal changes, many interactions before COVID-19 took place between students on lobbies, before and after the lectures and laboratories, and in breaks. Moreover, for the Algorithm Design course, students worked in laboratories, where they could discuss with teaching assistants in front of the computer screens. These interactions obviously required another channel of communication during online learning.

A second category of changes were probably not expected: The

metrics provided by the analysis with *ReaderBench* show that during COVID-19, the complexity indices of the exchanged texts increased. It is an interesting finding, which may be explained by the fact that, due to the lack of face-to-face conversations, where nonverbal communication also occurs, participants felt the need to be more communicative in the online exchanged messages.

The generated sociograms can be used by teachers to follow the evolution of students in term of interactions, interactivity, and online participation, thus enabling them to intervene when they notice a decrease in participation or inactivity. Students' chances of passing or obtaining a better grade can be increased if teachers encourage them to be more engaged throughout the course. As many learning institutions moved to online and face-to-face interactions and activities were drastically reduced, a mechanism that keeps track of students' activities and their likelihood of obtaining good scores would be beneficial for both students and teachers.

One potential limitation of this study regards the generalizability of the model. One factor to consider regards students who exhibited lurking behaviors and did not have any active posts. These students could not be included within the analyses as there were no textual traces to analyze. This factor may influence the accuracy of model predictions when considering the entire population of students. Thus, additional mechanisms centered on log analytics (e.g., access to specific posts, homework completion rates) need to be taken into account to build baseline models capable of generating predictions for this category of students. Moreover, the current experiments need to be applied on a larger timeframe and on additional courses in order to create generalizable models across different course topics and situations. The scope of this paper is limited to introducing the updated processing pipeline and the building blocks for creating custom prediction models, while the current in-depth analyses were conducted on a course in two consecutive yearly installments, before and during the pandemic. Longer timeframes (i.e., more installments) and additional courses will be considered to create predictive models with a higher degree of generalizability, which is hard, if not impossible, to obtain with the changes induced by the pandemic. Nevertheless, the specificities of each course, with underlying topics, activities, and interaction patterns, need to be considered while building prediction models.

In terms of future work, our aim is to introduce in-class evaluations in which we assess the impact of using the automated tools during the academic year that enable timely reactions from the tutors, in contrast to a posteriori assessments. In addition, we will introduce custom configurations that consider specific semantic models tailored for the topic of each course. This will afford applications beyond the course targeted in this study. Another overarching objective is to provide automated feedback to both students and teachers regarding student involvement in the course. To this end, the CNA plugin will be expanded such that Tutors will be able to periodically evaluate the evolution of their students, identify at-risk students, and take timely actions. Further modifications are envisioned wherein students can be incentivized to more actively engage and collaborate more with their peers. Ultimately, the objective is to maximize students' chances of success in the course.

Credit author statement

Maria-Dorinela Dascalu: Software, Visualization, Writing – original draft, Stefan Ruseti: Software, Formal analysis, Writing – original draft, Validation, Mihai Dascalu: Conceptualization, Methodology, Writing – original draft, Funding acquisition, Danielle S. McNamara: Methodology, Writing – review & editing, Funding acquisition, Mihai Carabas: Data curation, Investigation, Traian Rebedea: Conceptualization, Validation, Writing – review & editing, Stefan Trausan-Matu: Validation, Supervision, Writing – review & editing.

References

Achieve Virtual. (n.d.). Infographic: Evolution of virtual education. Retrieved September 1st, 2020, from https://achievevirtual.org/infographic-evolution-of-virtual-education/.

Allen, L. K., Jacovina, M. E., Dascalu, M., Roscoe, R., Kent, K., Likens, A., et al. (2016). {ENTER}ing the time series {SPACE}: Uncovering the writing process through keystroke analyses. In *9th int. Conf. On educational data mining (EDM 2016)* (pp. 22–29). Raleigh, NC: International Educational Data Mining Society.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(4–5), 993–1022.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems, 30,* 1–7.

Cao, W., Fang, Z., Hou, G., Han, M., Xu, X., Dong, J., et al. (2020). The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry Research,* 112934.

Chick, R. C., Clifton, G. T., Peace, K. M., Propper, B. W., Hale, D. F., Alseidi, A. A., et al. (2020). Using technology to maintain the education of residents during the COVID-19 pandemic. *Journal of Surgical Education, 77*(4), 729–732. https://www.sciencedirect.com/science/article/pii/S1931720420300842?casa_token=RedC882U7uIAAAAA:iCVBTGQFG6JjIaXCRSJibC2_GfpJayYLVI4vzAM0fl_e0uy1ug-gLG_If4i11PVPnVGkK-9-734.

Cioaca, V., Dascalu, M., & McNamara, D. S. (2020). Extractive summarization using Cohesion Network Analysis and submodular set functions. In *22nd international symposium on symbolic and numeric algorithms for scientific computing (SYNASC 2020).* Online: IEEE.

Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., et al. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Learning & Teaching, 3*(1), 1–20.

Crossley, S., McNamara, D. S., Baker, R., Wang, Y., Paquette, L., Barnes, T., & Bergner, Y. (2015). Language to Completion: Success in an Educational Data Mining Massive Open Online Class. *Proceedings of the 8th International Conference on Educational Data Mining.* International Educational Data Mining Society.

Crossley, S. A., Dascalu, M., Baker, M., McNamara, D. S., & Trausan-Matu, S. (2017). Predicting success in massive open online courses (MOOC) using Cohesion Network Analysis. In *12th int. Conf. On computer-supported collaborative learning (CSCL 2017)* (pp. 103–110). Philadelphia, PA: ISLS.

Crossley, S. A., Karumbaiah, S., Ocumpaugh, J., Labrum, M. J., & Baker, R. S. (2019). Predicting math success in an online tutoring system using language data and clickstream variables: A longitudinal analysis. In *2nd conference on language, data and knowledge (LDK 2019): Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.*

Crossley, S. A., Paquette, L., Dascalu, M., McNamara, D. S., & Baker, R. S. (2016). Combining click-stream data with NLP tools to better understand MOOC completion. In *6th int. Conf. On learning analytics & knowledge (LAK '16)* (pp. 6–14). Edingurgh, UK: ACM.

Dascalu, M. (2014). Analyzing discourse and text complexity for learning and collaborating. In *Studies in computational intelligence* (Vol. 534)Switzerland: Springer.

Dascalu, M., Crossley, S. A., McNamara, D. S., Dessus, P., & Trausan-Matu, S. (2018). Please ReaderBench this text: A multi-dimensional textual complexity assessment framework. In S. Craig (Ed.), *Tutoring and intelligent tutoring systems* (pp. 251–271). Hauppauge, NY, USA: Nova Science Publishers, Inc.

Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., & Nardy, A. (2014). Mining texts, learner productions and strategies with ReaderBench. In A. Peña-Ayala (Ed.), *Educational data mining: Applications and trends* (pp. 345–377). Cham, Switzerland: Springer.

Dascalu, M., McNamara, D. S., Trausan-Matu, S., & Allen, L. K. (2018). Cohesion Network Analysis of CSCL participation. *Behavior Research Methods, 50*(2), 604–619. https://doi.org/10.3758/s13428-017-0888-4

Dascalu, M.-D., Ruseti, S., Carabas, M., Dascalu, M., Trausan-Matu, S., & McNamara, D. S. (2020). Cohesion Network Analysis: Predicting course grades and generating sociograms for a Romanian Moodle course. In *16h int. Conf. On intelligent tutoring systems (ITS 2020).* Online: Springer.

Dascalu, M., Trausan-Matu, S., McNamara, D. S., & Dessus, P. (2015). ReaderBench – automated evaluation of collaboration based on cohesion and dialogism. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 395–423.

Dougiamas, M., & Taylor, P. (2003). Moodle: Using learning communities to create an open source course management system. In *World conference on educational multimedia, hypermedia and telecommunications (EDMEDIA) 2003. Chesapeake, VA, USA.*

Gers, F. A., Eck, D., & Schmidhuber, J. (2002). *Applying LSTM to time series predictable through time-window approaches Neural Nets WIRN Vietri-01* (pp. 193–200). Springer.

Hiltz, S. R., & Turoff, M. (2005). Education goes digital: The evolution of online learning and the revolution in higher education. *Communications of the ACM, 48*(10), 59–64.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*(2), 211–240.

Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education, 45*.

Masala, M., Ruseti, S., & Dascalu, M. (2020). RoBERT – a Romanian bert model. In *28th int. Conf. On computational linguistics (COLING)* (pp. 6626–6637). Barcelona, Spain (Online): ACL.

McNamara, D. S., Allen, L. K., Crossley, S. A., Dascalu, M., & Perret, C. A. (2017). Natural Language processing and learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 93–104). Society for Learning Analytics Research.

Mihalcea, R., & Tarau, P. (2004). TextRank: Bringing order into texts. In *Conference on empirical methods in Natural Language processing (EMNLP 2004)* (pp. 404–411). Barcelona, Spain: ACL.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representation in vector space. In *Workshop at ICLR*. AZ: Scottsdale.

Moodle Docs. (n.d.). About Moodle. Retrieved September 1st, 2020, from https://docs.moodle.org/39/en/About_Moodle.

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education, 14*(2), 129–135.

Nistor, N., Dascalu, M., Serafin, Y., & Trausan-Matu, S. (2018). Automated dialog analysis to predict blogger community response to newcomer inquiries. *Computers in Human Behavior, 89*, 349–354. https://doi.org/10.1016/j.chb.2018.08.034

Nistor, N., Dascalu, M., Tarnai, C., & Trausan-Matu, S. (2020). Predicting newcomer integration in online learning communities: Automated dialog assessment in blogger communities. *Computers in Human Behavior, 105*.

Nistor, N., Dascalu, M., & Trausan-Matu, S. (2016). Newcomer integration in online knowledge communities: Exploring the role of dialogic textual complexity. In *12th int. Conf. On learning sciences (ICLS 2016)* (pp. 914–917). Singapore: International Society of the Learning Sciences (ISLS).

Pappas, C. (2019). *The top open source learning management systems [2019 update]*. Retrieved September 1st, 2020, from https://elearningindustry.com/top-open-source-e-learning-management-systems.

Scott, J. (2017). *Social network analysis*. London, UK: Sage.

Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. E. (2014). Who does what in a massive open online course? *Communications of the ACM, 57*(4), 58–65.

Sharma, K., Jermann, P., & Dillenbourg, P. (2015). Identifying styles and paths toward success in MOOCs. In O. C. Santos, J. G. Boticario, C. Romero, M. Pechenizkiy, A. Merceron, P. Mitros, et al. (Eds.), *8th int. Conf. On educational data mining* (pp. 408–411). Madrid, Spain.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. *In 2018 17th IEEE international conference on machine learning and applications (ICMLA)* (pp. 1394–1401). IEEE.

Sirbu, M.-D., Dascalu, M., Crossley, S. A., McNamara, D. S., Barnes, T., Lynch, C. F., et al. (2018). Exploring online course sociograms using Cohesion Network Analysis. In C. P. Rosé, R. Martínez-Maldonado, U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, et al. (Eds.), *19th int. Conf. On artificial intelligence in education (AIED 2018), Part II* (pp. 337–342). London, UK: Springer.

Sirbu, M.-D., Dascalu, M., Crossley, S. A., McNamara, D. S., & Trausan-Matu, S. (2019). Longitudinal analysis of participation in online courses powered by Cohesion Network Analysis. In *13th int. Conf. On computer-supported collaborative learning (CSCL 2019)* (pp. 640–643). Lyon, France: ISLS.

Trausan-Matu, S. (2020). The polyphonic model of collaborative learning. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge international handbook of research on dialogic education* (pp. 454–468). New York, NY, USA: Routledge.

Trausan-Matu, S., & Slotta, J. D. (in press). Artifact analysis. In U. Cress, Oshima, C. P. Rose & A. Wise (Eds.), The international handbook of computer-supported collaborative learning. Cham, Switzerland: Springer.

Tu, C.-H. (2002). The measurement of social presence in an online learning environment. *International Journal on E-Learning, 1*(2), 34–45.

Weidlich, J., & Bastiaens, T. J. (2019). Designing sociable online learning environments and enhancing social presence: An affordance enrichment approach. *Computers & Education, 142*, 16.

Wen, M., Yang, D., & Rose, C. P. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *7th int. Conf. On educational data mining* (pp. 130–137). London, UK.

Yilmaz, F. G. K., & Yilmaz, R. (2020). The impact of feedback form on transactional distance and critical thinking skills in online discussions. *Innovations in Education & Teaching International, 57*(1), 119–130.