Preprints.org

Article

Not peer-reviewed version

Exploiting Properties of Student Networks to Enhance Learning in Distance Education

Rozita Tsoni * , Evgenia Paxinou , Aris Gkoulalas-Divanis , Dimitrios Karapiperis , Dimitrios Kalles , Vassilios S. S. Verykios

Posted Date: 11 March 2024

doi: 10.20944/preprints202403.0638.v1

Keywords: distance learning; learning analytics; social network analysis

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Exploiting Properties of Student Networks to Enhance Learning in Distance Education

Rozita Tsoni [1,*], Evgenia Paxinou [1], Aris Gkoulalas-Divanis [2], Dimitrios Karapiperis [3], Dimitrios Kalles [1] and Vassilios S. Verykios [1]

[1] Hellenic Open University, Patras, Greece
[2] Merative, Healthcare, Dublin, Ireland
[3] International Hellenic University, Thermi, Greece
* Correspondence: rozita.tsoni@ac.eap.gr

Abstract: Distance Learning has become the new standard, especially after the pandemic and due to the technological advances, that are incorporated into the teaching procedure. At the same time, the augmented use of the internet has blurred the borders between distance and conventional learning. Students interact mainly through LMSs, leaving their digital traces that can be leveraged to improve the educational process. This work aims to propose an analysis model that can capture the students' behaviors based on the clickstream data associated with the discussion forum and additionally to suggest interpretable patterns that will support education administrators and tutors in the decision-making process. To achieve our goal, we use Social Network Analysis (SNA) as networks represent complex interactions in a meaningful and easily interpretable way. Moreover, simple or complex network metrics are becoming available to provide valuable insights into the students' social interaction. This study concludes that by leveraging the imprint of these actions in an LMS and using metrics of SNA, differences can be spotted in the communicational patterns that go beyond simple participation recording. Although HITS and PageRank algorithms were created with completely different targeting, it is shown that they can also reveal methodological features in students' communicational approach.

Keywords: distance learning; learning analytics; social network analysis

1. Introduction

Distance Learning (DL) appeared over a century ago as a modern and innovative method in education. A robust theoretical framework has been created, which is still evolving. DL has become the new standard, especially after the pandemic and due to the technological advances, that are incorporated into the teaching procedure. At the same time, the augmented use of the internet has blurred the borders between distance and conventional learning. The Learning Management System (LMS) was first introduced in the late 1990s (Davis, Carmean & Wagner, 2009) to provide instructors with a way to develop and deliver their educational material, observe their students' participation, and assess their performance. An LMS aims at replacing the conventional classroom, by constituting a central setting where learning occurs.

In DL, more than any other educational method, the teaching and learning process is efficient if there is constant communication and interaction between those who are involved (Simpson, 2000). DL, as a teaching and learning strategy, may have an inherent disadvantage: learners who attend DE programs are physically separated from their tutors and peers (Panagiotakopoulos et al., 2013). Thus, an important additional goal of DL is to enhance students' autonomy. Self-regulated learning was strongly associated with acquiring knowledge and skills by becoming aware of the appropriate strategies and having the ability to use them effectively (Zimmerman, 1990). Having high levels of metacognition, having *"the ability to control one's cognitive processes"* (Livingston, 2003), is also a characteristic of a learner with critical awareness. Undoubtedly, there are a lot of different learning

© 2024 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

paths leading to effective learning. The available technological tools and the educational designing process play a pivotal role in overcoming obstacles, like distance and timing. Miyazoe & Anderson (2013) introduced *the "Equivalency Theorem"* which posits that:

1. "Deep and meaningful formal learning is supported, as long as one of the three forms of interaction (student-teacher; student-student; student-content) is at a high level. The other two may be offered at minimal levels, or even eliminated, without degrading the educational experience.
2. High levels of more than one of the above three modes will likely provide a more satisfying educational experience, although these experiences may not be as cost- or time-effective as less interactive learning sequences."

Moreover, distance learning adult students are struggling to combine studying and educational tasks with family and work obligations, during the working days. Therefore, they log in to the institutional LMS to communicate through fora with their peers and their tutors, mostly during evenings and weekends (Kagklis et al., 2015). Therefore, tutors try to be present and supportive of their students, in a minimum time pan. By monitoring their students' participation in the LMS discussion fora, instructors realize that it is of utmost importance to model the learners' behavioral patterns in these environments (Geigle & Zhai, 2017).

Learning analytics (LA) can provide the information on the students' behavior, that tutors need to have for assisting them in their self-directed learning procedure. At the same time, students can preserve their privilege to study in their place, at their own pace without having to be physically present on a campus. Empirical findings from a trans-European study (Wollny et al., 2023) indicate a high demand for LA and a certain lack of confidence in meeting the high expectations that the educational community has set for the benefits that LA can offer. The process of capturing complex students' interactions in an educational environment, is far from simple. This challenge can be approached by doing small steps, each time, aiming at specific features. According to Setiawan et al. (2020), when students are enrolled in an online course, it is feasible to mine a large amount of data from the platform logins, allowing the detection and processing of the behavioral logs. Modeling is a helpful way to automatically capture students' interactions, in a course discussion forum. In DL where most of the learning occurs in unsupervised environments, extracting and analyzing large amounts of forum data, could lead to deriving useful knowledge and improving the design of a course.

This study aims to identify students' behavior patterns, through their logging into the discussion forum of a DL module, at the Hellenic Open University (HOU), as an attempt to identify different learning approaches in DE. In the discussion forum, students log in and address a query, reply to a peer's question, participate in a discussion thread, or just check on the latest posts. Our goal is firstly, to design a model that may capture the aforementioned students' actions (behaviors) based on the clickstream data associated with the discussion forum and secondly, to suggest interpretable patterns that will support education administrators and tutors in the decision-making process. To achieve our goal, we use Social Network Analysis (SNA) as networks represent complex interactions in a meaningful and easily interpretable way. Additionally, simple or complex network metrics are available to provide valuable insights into the students' social interaction. An additional, yet not less important, goal is to highlight the differences between network metrics interpretation and the knowledge that they can provide concerning students' behavior. Given that these metrics are by definition highly correlated, usually they are considered as similar and they are not interpreted separately in the relevant context. Here, we attempt to highlight their different meaning and the additional information that adds up while using Social Network Analysis in an educational context.

2. Related Work

LA is the process of converting raw data into meaningful knowledge, regarding learning. LA methodology mainly aims to understand and optimize the learning processes and also to improve the environments in which these processes occur (Siemens & Baker, 2012). At DE, discussion fora enable communication between students and instructors and therefore play a central role in learning,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

as they provide satisfaction and they enhance motivation and knowledge retention (Brindley, Walti & Blaschke, 2009; Tsoni et al., 2019). During the online learning, many data are recorded and accumulated in the institutional LMSs (Motz, Quick, Wernert & Miles, 2021). These data not only present the students' effort and behavior in a holistic way, but they also lead to very important outcomes, if they are interpreted by LA techniques (Lang et al., 2017; Tsoni et al., 2022; Tsoni, Panagiotakopoulos & Verykios, 2021). These interpretations can be used in the wider framework that could include concepts, such as *the community of practice* or *student-centered learning*, in an attempt to enhance teaching and learning. As social interaction has long been established as a major factor that also affects learning, SNA fits the criteria for imprinting communication and learning patterns. Lee et al. (2018) studied the students' preferences, while, i.e. they were watching educational videos, and used the networks formed between them, to extract behavioral patterns. Additionally, Sturludottir et al. (2021), found strong similarities between the networks created by students with the same course choices, and their actual major specialization in the latter studies. The changes, a network of a forum community may undergo during an academic year, were studied by Tsoni et al. (2020) and Lopez-Flores et al. (2022). These two types of research showed significant changes in graph density (that measures the number of ties between the nodes) and participation. Students' outdegree and network cohesion metrics are also identified as predictors of successfully completing the studies.

Simple metrics, like indegree and outdegree, provide useful information about students' participation in a forum community. However, Huang et al. (2014) claim that "superposting" does not necessarily imply a qualitative contribution to the forum community. The idea of finding centrality metrics to evaluate the contribution of those who post in a discussion forum, came from studies where researchers develop iterate algorithms, such as the PageRank algorithm, to calculate influence weights for citing articles based on the number of times that they have been cited (Pinsky & Narin, 1976; Brin & Page, 1998; Davis, 2008). Sanchez et al. (2021) highlighted the use of eigenvector centrality, as an indicator of the students' academic performance in the pilot course of mathematics. Additionally, several SNA metrics were positively strongly correlated with academic performance metrics (Putnic et al., 2016). However, it has to be noted that in all of the above studies, participating in the forum was a part of organized activities, embedded in the curriculum. Thus, participation was compulsory and students were given external motives through grading, to interact via the forum.

The research conducted by Da Silva et al. (2019) revealed that engagement within the forum community was more pronounced during graded activities. Additionally, when this motivational factor was absent, communication experienced a reduction. The potential application of Social Network Analysis (SNA) metrics as indicators of academic performance is exemplified in the study by Hernández-García et al. (2015). In their work, Hernández-García et al. (2016) employed Gephi to create multiple visualizations capturing students' interactions. However, they also underscored the challenge of interpreting intricate metrics, especially for individuals lacking expertise in the field, despite the numerous possibilities offered by Gephi and related tools. In the research conducted by Adraoui et al. (2017), the Pajek program package was utilized, focusing on centrality metrics as predictors of academic performance.

Elaborated algorithms, used in SNA, can also shed light on educational research. The algorithms HITS and PageRank were initially introduced focusing on ranking webpages. They can capture the added value of a node due to its ties with nodes of high importance. HITS and PageRank quickly found use in a wide area of research including educational research. According to Google the underlying assumption in the PageRank algorithm is that the most known and valid websites are likely to receive more links from others (Chonyy, 2021). Jon Kleinberg developed the HITS algorithm, which is based on the Principle of Repeated Improvement, as the PageRank algorithm. Kleingeld (1999) introduced the "authority" and the "hub" metrics to rank pages on the Web. Two scores are assigned for each web page: its authority, which estimates the quality of the content of the page, and its hub, which estimates the quality of its links to other web pages. There are several studies using more complex SNA metrics. However, eigenvector centrality, PageRank and HITS algorithm, are less used in SNA studies than simpler metrics, like, degrees, closeness and betweenness centralities, even

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

though they were strongly positively correlated with academic performance metrics according to the meta-analysis of Saqr et al. (2022).

3. Methodology

In this study, we propose a simple model to represent the behavioral patterns derived from a discussion forum, in the portal of the HOU, a university that is advocating DE.

3.1. Participants

The participants are students enrolled in two annual courses, in a postgraduate DE program: course A and course B. The forum community of course A includes 16 students and their tutors, and the forum community of course B includes 23 students and their tutors. For privacy-preserving purposes, the students' and tutors' names are replaced by randomly generated pseudonyms. For example, Ast5 denotes a student enrolled in course A and Bt2 denotes a tutor in course B. Each course's forum represents a unique microcosm of student interaction, influenced by specific course content, structure, and participant dynamics. We chose not to aggregate these data sets in our methodological approach since this decision could obscure these nuanced differences, thereby diluting the specificity and relevance of our findings.

3.2. Dataset

In this study, we visualize behavior patterns as graphs, where a node represents a participant (student or tutor) and a directed edge indicates a reply, given from one participant to another. The HOU portal is hosted on the Moodle (Modular Object-Oriented Dynamic Learning Environment) platform. Thus, the data are retrieved as a Moodle log file, which contains the participants' actions in the fora. The pre-processing for the creation of a unipartite-directed graph, mainly consists of the following steps:

1. The actions with the indication "discussion created" and "post created", are separately assorted from the log file.
2. The "discussion created" actions provide information on the creation of new discussion threads. Each thread is assigned to the participant who created it (student or tutor).
3. Each post is assigned to the participant who uploaded it and to the corresponding discussion thread that belongs to.
4. Each participant is represented as a node.
5. An incoming edge to a node represents a reply to a discussion thread, this participant has created (i.e., if Ast5 has three incoming edges then that means that three participants had posted in the threads that Ast5 has created).
6. An outgoing edge of a node denotes the posts that this specific participant made to other participants' threads (i.e., if Bst2 has 8 outgoing edges, then that means that Bst2 had replied in the threads that 8 other participants had created).
7. A self-loop denotes that the participant who made a post and created a thread replied to his/her original post.

3.3. Metrics and algorithms

To understand the outcomes of this study, it is essential to give some information on the basic network metrics (In-degree, Out-degree, Degree, weighted in-Degree Weighted Outdegree, Weighted degree, Closeness centrality, Harmonic closeness centrality, Betweenness centrality, Eccentricity and Eigenvector centrality) and the algorithms (HITS and PageRank) used in the modeling conducted in this study. Herein there is a succinct description delineating the Social Network Analysis (SNA) metrics employed within the scope of this investigation.

The *Indegree* of a node represents the number of the participants that reply at the threads of a certain person. The *Outdegree* of a node indicates the number of the participants who have created the threads that this node (person) posts in. The *Degree* is the sum of the Indegree and the Outdegree.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

The *Weighted in-Degree* shows the number of replies that a participant has received in her/his threads. The *Weighted Outdegree* denotes the number of posts that a participant has made.

The abovementioned information sets the ground to introduce the following centrality measures. *Closeness Centrality* is based on the mean geodesic distance, that is the number of edges of the shortest path between two nodes. Knowing that every node condenses all its discussion threads and every edge condenses all the replies to the threads of this node, we expect short geodesic distances in our networks, and therefore, high values of closeness centralities. Additionally, *Eccentricity* represents the maximum distance over all the nodes of the network. We expect to have low values, due to the small size of the network. *Betweenness Centrality* is a measure that has an added value, concerning communication in the educational forum, showing a node's ability to connect other nodes. In an educational environment, we expect to see participants with high betweenness centrality who act as communication facilitators. They enhance students' engagement and increase the closeness centrality of peripheral participants, as they bridge nodes that otherwise would have been disconnected. In a directed network, *Eigenvector Centrality* captures the importance and the prestige, a node has. It is proportional to the sum of the centralities of the nodes that are straight-linked to it. Therefore, a node's eigenvector centrality mainly depends on its neighbours' characteristics. However, it has to be highlighted that zero indegree results in zero eigenvector centrality. Indeed, a node with an indegree equal to zero is a participant who did not receive any answer in all of his/her threads.

Advanced metrics of higher complexity are derived from elevated algorithms illustrating a node's value in a network, by the quality of its neighbors and the strength of their ties. *HITS algorithm* uses the metrics "Authority" and "Hub". It is a link analysis algorithm that was first developed by Jon Kleinberg (Kleinberg et al., 1999) in an attempt to rate the quality and the reliability of Web pages when the Internet was originally forming. Initially, a hub and an authority value are assigned in each node according to its incoming and outgoing edges. An iterative process begins correcting these values until a default point of convergence is met. A high value of hub means that the node points to high authorities i.e., nodes with valuable information, represented as nodes with high in-degree in a directed network. Respectively, a node with high authority is being pointed by good hubs in a mutually reinforcing relationship. A good hub adds value to an authority and subsequently, the authority becomes better, adding more value to the hub in a recurrent process that, after several iterations, converges to a final result.

A second relevant algorithm is the PageRank algorithm, which was initially designed as a measure of influence and was implemented by directed graphs. The PageRank score is calculated by initially assigning a numerical weight to each node and recalculating this weight by taking into account the number of ties of the connected nodes. PageRank as well as HITS are based on the Principle of Repeated Improvement which is an iterative process where an initial value is assigned to a node and then a re-weighting process begins re-assigning new values according to each node's connections until the convergence criteria are met.

The directed network, that is created, aims to represent behavioral features of human communication. Every piece of information derived from this interaction can make a difference and reveal details that might be crucial for understanding the learning profiles. The metrics of the HITS and PageRank algorithms clearly distinguish the difference of the impact of an incoming and an outgoing edge, facilitating the interpretation of the results. In a communication network, the process of repeated improvement, that these algorithms use, allows us to efficiently imprint the augmented influence of a person in the community as they establish their relations with other participants, by considering their level of influence. The biggest difference between PageRank and HITS algorithms is that HITS calculates the quality based on the hubness and authority value, while PageRank calculates the ranks based on the proportional rank passed around the sites (Chonny, 2021).

Additionally, we used students' grades to capture their academic performance and relate it with the features of their communication deriving from the SNA metrics. In Course A students had to hand on four written assignments so we used the variables WA1, WA2, WA3, WA4 and the Average

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

grade (Av. WA). In Course B there were three written assignments leading us to use the variables WA1, WA2, WA3 and the Av. WA respectively.

4. Results and Discussion

Digging into communication communities to reveal behavioral patterns, constitutes a multifactorial and complicated research problem. Typical visualizations can only depict a limited amount of information. On the other hand, network graphs are visualizations that offer an information-rich image, where complicated interactions are illustrated in a comprehensible way. Borgatti & Halgin (2011) highlighted the importance of the position of a node per se, for defining its properties. This means, that in every network the position of each node can capture features that would otherwise be difficult or confusing to describe. Furthermore, the network representation facilitates the computation of Social Network Analysis (SNA) metrics, which unveil characteristics that may not be readily apparent from the graphical depictions. In the subsequent tables (Table 1 and Table 2), a summary of descriptive statistics is provided for the variables utilized in Course A and Course B, respectively. This summary includes the minimum and maximum values, mean and standard deviation, as well as variance, skewness, kurtosis, and overall sum for each metric.

Table 1. Summary measures for Course A.

Course A								
Variable	Min	Max	Mean	Std. deviation	Varia nce	Skew ness	Kurt osis	Overall sum
WA1	7,5	10	9,83	0,65	0,42	-3,87	15,00	147,50
WA2	7	10	9,67	0,84	0,70	-2,82	7,94	145,00
WA3	7,5	10	9,47	0,81	0,66	-1,49	1,40	142,00
WA4	0	10	8,39	3,44	11,80	-2,32	4,09	125,80
Av. WA	6,75	10	9,34	1,03	1,06	-1,87	2,66	140,08
In-degree	0	4	1,27	1,33	1,78	0,69	-0,64	19,00
Out-degree	0	2	0,67	0,62	0,38	0,31	-0,40	10,00
Degree	1	4	1,93	1,10	1,21	0,89	-0,44	29,00
Weighted in-degree	0	6	1,73	2,09	4,35	1,06	-0,19	26,00
Weighted out-degree	0	3	0,87	0,92	0,84	0,94	0,52	13,00
Weighted degree	1	9	2,60	2,47	6,11	1,81	2,50	39,00
Eccentricity	0	4	0,87	1,19	1,41	1,47	2,09	13,00
Closeness centrality	0	1	0,34	0,41	0,17	0,67	-1,22	5,12
Harmonic closeness centrality	0	1	0,36	0,42	0,18	0,54	-1,48	5,36
Betweenness centrality	0	0,02	0,00	0,00	0,00	3,87	15,00	0,02
Authority	0	0,65	0,16	0,21	0,04	1,20	0,47	2,44
Hub	0	0,27	0,03	0,07	0,01	3,10	10,03	0,42

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

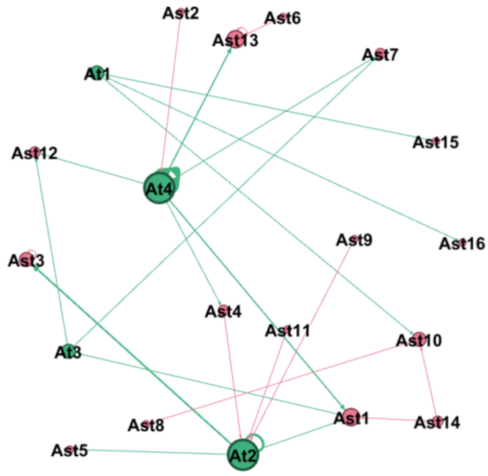
PageRank	0,02	0,06	0,03	0,01	0,00	1,01	0,06	0,46
Eigenvector Centrality	0	1	0,19	0,29	0,09	1,83	3,16	2,85

Table 2. Summary measures for Course B.

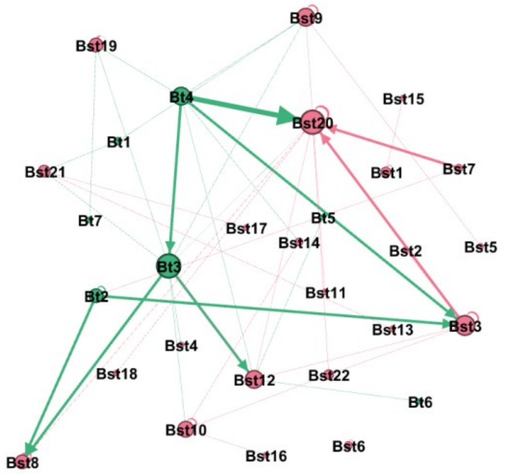
Course B								
Variable	Min	Max	Mean	Std. deviation	Variance	Skewness	Kurtosis	Overall sum
WA1	5	10	8,22	1,63	2,64	-1,09	0,13	180,90
WA2	0	10	7,35	2,65	7,02	-1,45	1,62	161,70
WA3	0	10	7,50	3,04	9,24	-1,61	1,69	165,00
Av. WA	2,9	9,7	7,69	2,11	4,47	-1,21	0,52	169,20
In-degree	0	9	2,09	2,64	6,94	1,15	0,53	46,00
Out-degree	0	3	1,23	0,75	0,56	1,07	1,56	27,00
Degree	1	10	3,32	2,77	7,66	1,02	-0,04	73,00
Weighted in-degree	0	13	2,64	3,54	12,53	1,46	1,95	58,00
Weighted out-degree	0	4	1,45	1,06	1,12	1,06	0,30	32,00
Weighted degree	1	14	4,09	3,95	15,61	1,24	0,54	90,00
Eccentricity	0	2	0,77	0,69	0,47	0,32	-0,70	17,00
Closeness Centrality	0	1	0,59	0,47	0,22	-0,43	-1,83	12,93
Harmonic Closeness Centrality	0	1	0,60	0,47	0,22	-0,49	-1,81	13,17
Betweenness Centrality	0	0,03	0,00	0,01	0,00	4,64	21,64	0,03
Authority	0	0,57	0,13	0,17	0,03	1,11	0,54	2,83
Hub	0	0,29	0,07	0,09	0,01	1,26	1,13	1,64
PageRank	0,01	0,04	0,01	0,01	0,00	1,89	3,99	0,27
Eigenvector Centrality	0	1	0,11	0,22	0,05	3,30	12,56	2,50

To leverage the abovementioned benefits, we created two directed unipartite networks for courses A and B, shown in Figure 1. Each node represents a forum participant who could be a tutor (green node) or a student (pink node). The magnitude of the nodes is proportional to their degree. Thus, large nodes represent participants who posted a lot and received many replies. The edges are colored according to the origin node, showing that the post was submitted by a student or a tutor, and their width is proportional to their weight, which is the number of posts. In some nodes, the small, semicircular lines represent self-loops, which is a connection of a node with itself and visualizes a participant's reply to this own thread.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1



(a)



(b)

Figure 1. Networks formed based on the participants' communication, through the discussion Forum, in (a) course A and (b) course B.

In both networks, tutors' contribution is clear. Tutors seem to be the leaders in the network interactions. They have a binding role in the community, acting as communication facilitators (a tutor's main responsibility in DE). The average path length in course A is 1.643 and 1.608 in Course B, indicating that the average distance between to random nodes is approximately the same in both networks. The network diameter is equal to 4 for Course A and equal to 3 for Course B. Therefore, it takes 4 hops to travel across the most distance nodes in the first Course, while in Course B it takes 3 hops. The average path length in course B is 1.608 and the network diameter is, despite the larger participation compared to course A.
In course A, the connections in communication are simpler than in course B: Students tend to reach their tutors for i.e. posing a question, rather than their peers. This is an indication to the community that the trust and collaboration between peers are still at a premature level, as they prefer to interact with the "expert" who is for them "the more knowledgeable other" (Vygosky, 1987). However, according to Figure 1, some participants have an equally important role in the network, as their tutors'. To thoroughly examine this role and identify different approaches to learning between students, we commuted the Social Network Analysis (SNA) of these metrics presented in Section 3. The overall participation is mainly captured by the total weighted degree. The weighted out-degree

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

shows the tendency to participate in other participants' discussions and the in-degree shows the interest that creates a participant's posts.

In course A, students Ast13 and Ast3 have the two highest weighted degrees, weighted in-degrees, weighed-out-degrees, PageRank scores, and Eigenvector centralities. Interestingly, both students Ast3 and Ast13 (Figure 1a) owe their beneficial position to their connections with tutors. Student Ast3 is connected exclusively with his/her tutor (Figure 2). Additional value to his/her eigenvalue centrality is added by the self-loops, that is, the replies he/she makes in his/her threads. That means that the student continues to participate at the dialogue that she/he started, commending at the answer of a co-learner or a tutor posted at her/his thread. This behavior leads the students to gain an accumulative advantage due to the Matthew effect (the tendency to accrue social success in proportion to their initial level of popularity and number of friends) (Rigney, 2010) in means of importance in the communication network.

Student Ast1 is also very active receiving many replies in the discussions that he/she created. For student Ast1 the weighted outdegree is zero, meaning that she/he did not reply in any of her/his peers discussions. She/he only participated in discussions created by her/himself. In the contrary, student Ast14 replied many times in other participants' threads although she/he did not start any conversation. Therefore, he/she obtains a high hub score in the network, along with Ast6 and Ast8. Although the latter two students are not very active, they reply in threads created by influential participants (high authority scores), gaining importance. The best authorities scores of the network belong to the nodes Ast1, Ast12 and Ast7 (see Appendix). Except for Ast1, the other two are not the most popular nodes, in means of the number of replies received. However, they also gain credit by attracting replies from prestigious participants who make them the best authorities.

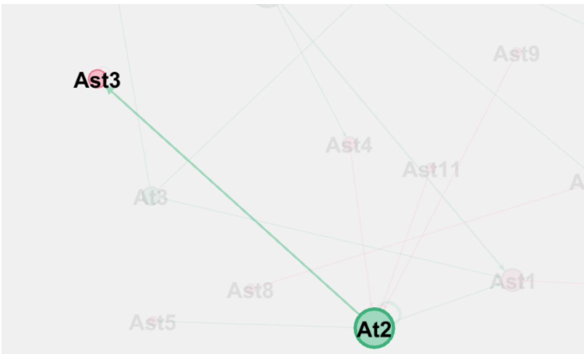


Figure 2. The "exclusive" communication of Ast3 with his/her tutor.

The node Ast4 is not included in any of the top three rankings of importance measures (Authority, Hub, PageRank, Eigenvector) and most of its metrics values are relatively low. However, it plays an important role in the communication network. It is the only node that has non-zero betweenness centrality, actively contributing to bridging the gap between two disconnected areas of the network.

In course B, Bst20, Bst3 and Bst8 own the most popular posts. Students Bst20 and Bst3 are also in the top three best authorities. Yet, Bst9 has higher authority in the HITS algorithm, compared to Bst8. That is because they received more replies, made by participants with higher influence (Figures 3 and 4).

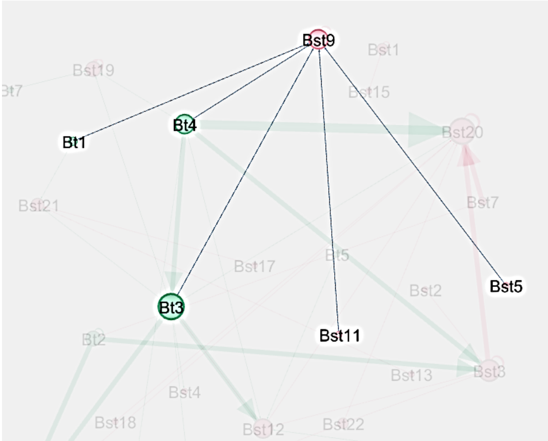


Figure 3. Bst9's connections.

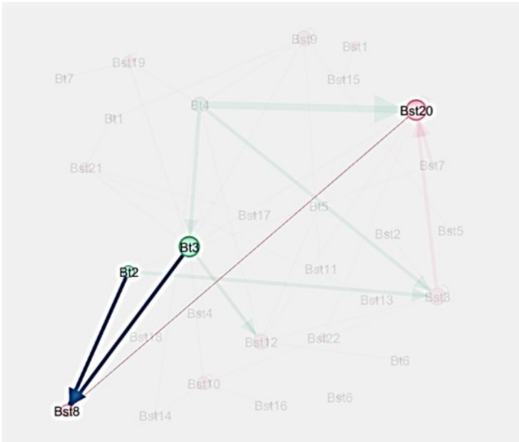


Figure 4. Bst8's connections.

Concerning the participation in other discussions, the most active students were Bst12, Bst3, and Bst8 (higher weighted degree). However, the best hub scores were encountered in nodes Bst22, Bst12 and Bst7. This is mainly due to their multiple connections with Bst20, which is one of the most important nodes of the network (ranking first in Weighted In-Degree, Weighted Degree, Authority, PageRank and Eigenvector Centrality). There is a totally different story concerning students' mediative role. The top three students regarding betweenness centrality, were Bst12, Bst3, Bst14 and Bst8. The "star" student Bst20 presents zero betweenness centrality. This situation reflects a different learning approach. While Bst3 and Bst8 are actively participating, creating popular discussion threads, and replying to other discussions, even from peripheral participants, acting as a bridge, Bst20 rarely replies, but he/she created threads where important participants post, gaining influence, only participating in his/her posts. Student Bst3 is also a notable case since he/she is included in the top three of Weighted Degree, Authority, PageRank, Betweenness and Eigenvector Centrality rankings. His/her actions are also targeted however, he/she is more outgoing, replying to his/her peers, even if their post is not popular, showing collaborative spirit.

As it was shown, different metrics reveal a different aspect of each participant's contribution to the discussion community. Each student is represented by a different combination of metrics values that can be shown graphically. To visualize the differences between students' SNA metrics, in a common graph, we applied a min-max normalization (minimum=0, maximum=1). The results were reported in a heatmap (Figures 5 and 6) where dark blue represents 0, white represents 0.5 and dark red represents 1.

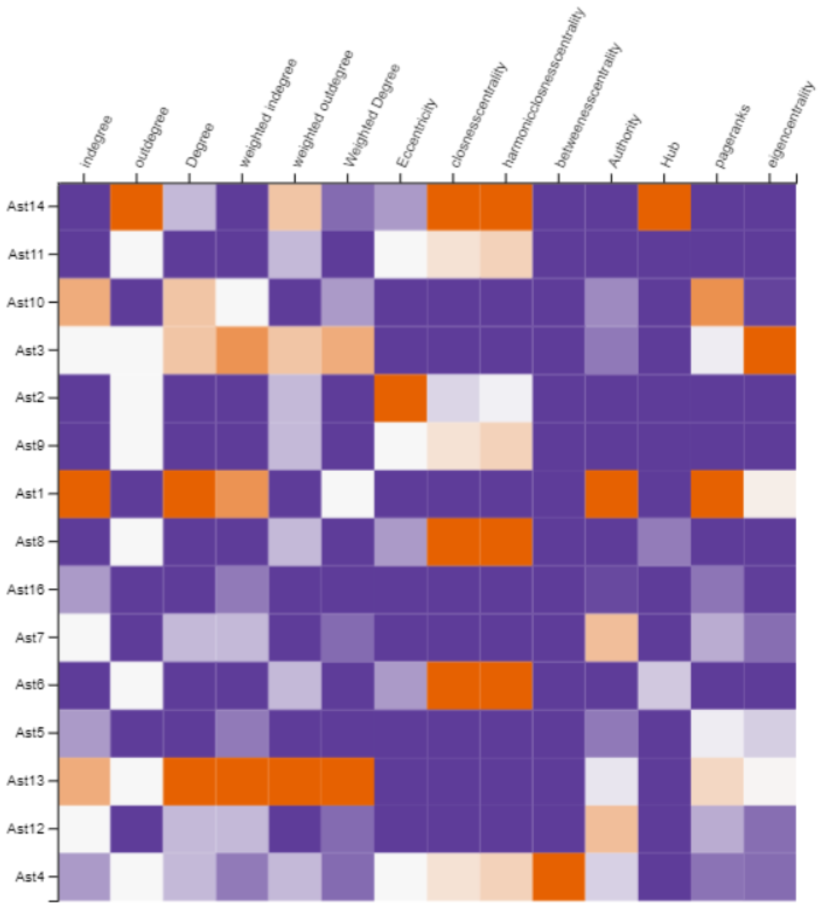


Figure 5. Students' SNA metrics for course A.

Figure 5 can be seen as a condensed profiling graph where different communication approaches are becoming obvious. For example, let's study students Ast8 and Ast13. Ast8 has a low number of posts and replies, but due to certain interactions, he/she is in the center of the network (high closeness centrality), while Ast13 is active but peripheral.

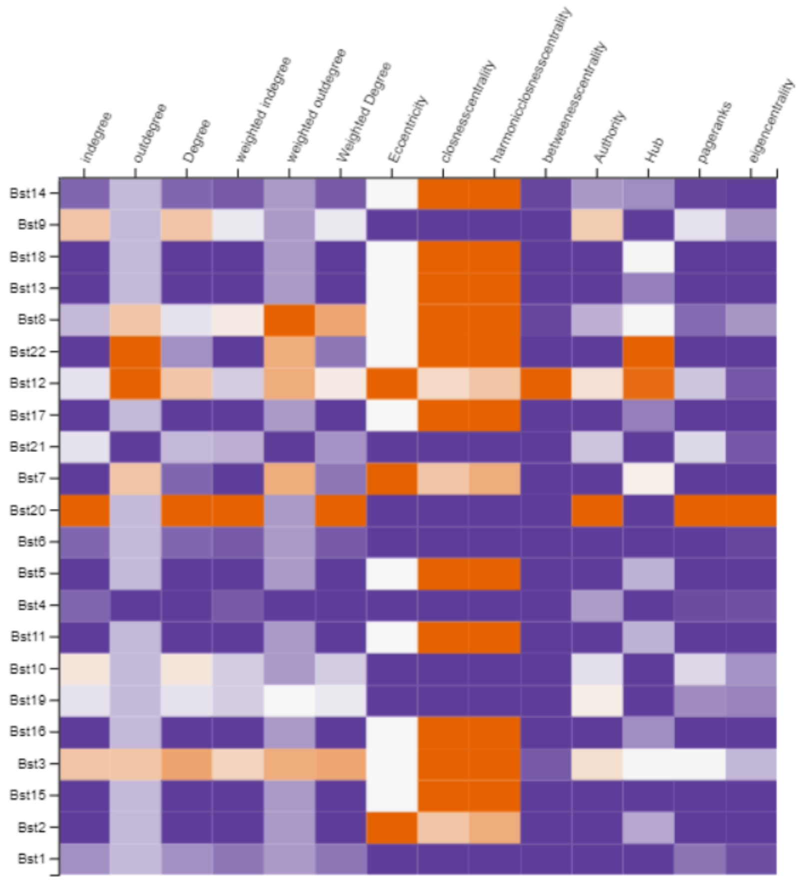


Figure 6. Students' SNA metrics for course B.

Similarly, in Figure 6 different behaviors can also be spotted. Bst3 represents a very active student, with a central role in the network. Instead, Bst1 is one of the most isolated students with low participation, in a less prestigious position.

Previous research (Tsoni et al., 2021, 2022) has shown that three important factors affect learning: online participation, academic achievement and position in the communication network. It was therefore considered useful to examine the relationship between SNA metrics and academic performance. The attributes WA1, WA2, WA3, WA4 and mean WA, represent the grades in four written assignments (WA) and their mean value, correspondingly. A correlation analysis was conducted for both courses. The majority of correlations between grades and Social Network Analysis (SNA) metrics were found to be statistically insignificant. This is likely attributed to the varied usage patterns of the forum within these courses. Participation was voluntary, there were not any mandatory learning activities within the forum, and students utilized it for diverse purposes: connecting with peers, posing queries related to the course material, receiving updates on deadlines and grades, or simply socializing. Nonetheless, certain statistically significant correlations were observed and are detailed below. Tables 3 and 4 present the variables that exhibited statistically significant correlations, along with their correlation values and corresponding p-values. Given our focus on exploring the relationship between forum participation and academic performance, only such correlations have been included in these tables.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Table 3. Statistically significant correlation, relation SNA metrics and grades for Course A.

Course A			
Variable A	**Variable B**	**Correlation value**	**p value**
WA1	Eccentricity	-0,730	0,002*
WA2	Out-degree	-0,644	0,010*
WA2	Hub	-0,788	0,000*
WA3	Weighted outdegree	-0,583	0,023*

Table 4. Statistically significant correlation, relation SNA metrics and grades for Course B.

Course B			
Variable A	**Variable B**	**Correlation value**	**p value**
WA1	PageRank	-0,448	0,037*
WA1	Eigenvector centrality	-0,513	0,015*
WA2	PageRank	-0,433	0,044*
WA2	Eigenvector centrality	-0,432	0,045*

Because of the extensive array of metrics utilized in this study, the correlation matrix may prove challenging to interpret. Graphs were used as a means to visually summarize complex data sets succinctly. This method was chosen to facilitate a more accessible understanding of patterns across a broad audience, including those who may not specialize in quantitative analysis. Consequently, an alternative presentation method was adopted. The correlation matrix was rendered as a heatmap, wherein the correlation coefficient was depicted using a color scheme (with -1 indicated by red and +1 by blue), and the outcomes are displayed in Figure 7 and Figure 8.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

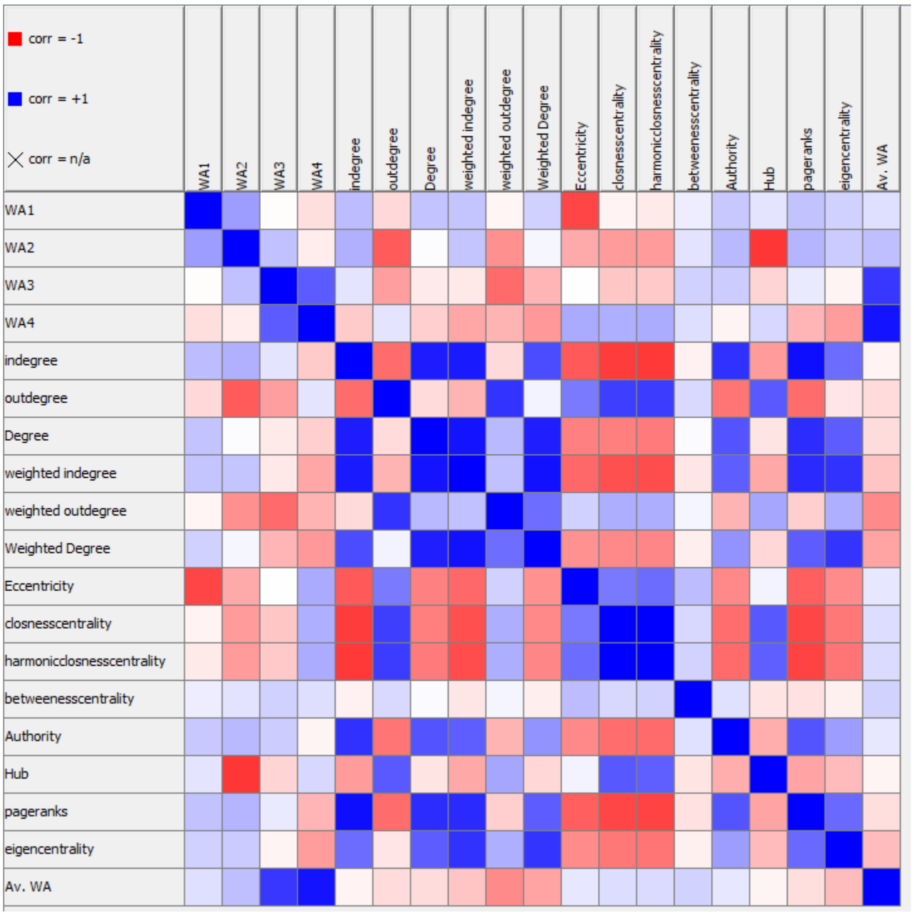


Figure 7. The correlation matrix between grades and SNA metrics for course A.

In course A (Table 3), there is a strong negative correlation between the grade of the first written assignment (WA1) and Eccentricity (r (13) =-.73, p<0.005) and a moderate negative correlation between the grade of the second written assignment (WA2) and Outdegree (r (13) =-.64, p<0.01). Additionally, there is a moderate negative correlation between the grade of the third written assignment (WA3) and Weighted Outdegree (r (13) =-.58, p<0.05). The negative correlation may reflect the need of certain students to communicate and discuss the difficulties they encounter. High SNA metrics, along with low grades, correspond to students who seek answers to their questions through forum communication. This suggestion is also supported by the structure of the network, where tutors act as communication facilitators providing students with answers.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

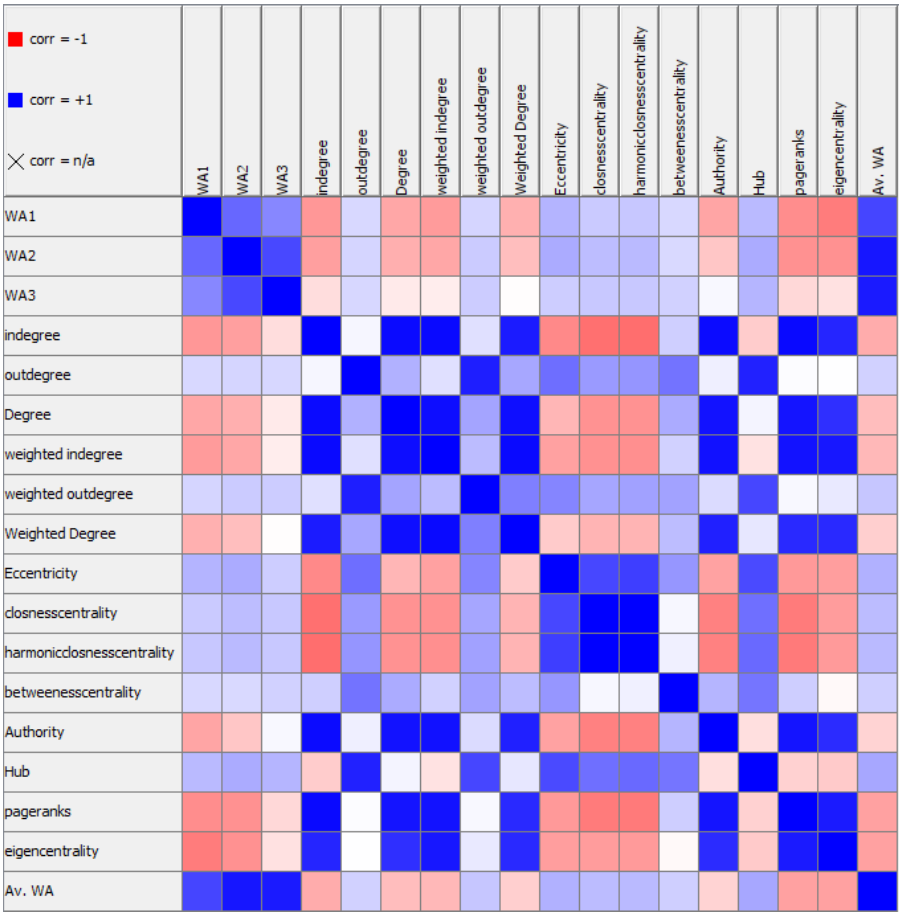


Figure 8. The correlation matrix between grades and SNA metrics for Course B.

Similar results are presented in Course B (Table 4). There is a moderate negative correlation between the grade of the first written assignment (WA1) and Eigenvector Centrality (r (20) =-.51, p<0.05) and a weak negative correlation between the grade of the first written assignment (WA1) and PageRank score (r (20) =-.45, p<0.05). There is also a weak negative correlation between the grade of the third written assignment (WA3) and PageRank score (r (20) =-.43, p<0.05) and between the grade of the third written assignment (WA3) and Eigenvector Centrality (r (20) =-.43, p<0.05). Other strong correlations appearing in the graph are either irrelevant, capturing the structural affinity of network metrics, or not statistically significant (p>0.05). The majority of the studies in the literature, that correlate SNA metrics with academic performance, found positive correlations between them (Saqr et al., 2022). However, as it is aforementioned, the SNA metrics are derived from forum activities that are a part of the students' workload. In such cases, positive correlations are expected, since it is expected for diligent students to have good grades. Kipling et al. (2023), in their recent work, present a critical view of the effectiveness of providing external motives for forum use. More specifically, it is stated that certain attempts to control engagement *"may be proven particularly ineffective stimulating unhelpful grade-focused participation"*. In general, when forum activities are structured and graded, there is external motivation for the students to participate. Thus, forum activity becomes another assignment for them. Measuring forum participation in such cases is, in fact, equivalent to capturing one more grade. In this work, we analyze forum participation as an indication of genuine and optional interaction. This means that forum participation metrics capture students' social interaction and collaboration patterns, reflecting on their learning behavior within a group of peers. The results showed that the students use the forum to address their difficulties and solve their course-related problems. This is a plausible explanation of the negative correlations, showing that the bigger the barriers they face, the more they pose questions and interact with their tutors and peers.

5. Conclusions

Communication, interaction and dialogue are important concepts of distance education. Already from the early '80s, Holmberg (1983) introduced the theory of "Guided didactic conversation" which suggests that autonomous learning in a learner-centered open environment is promoted through constant communication between "the educans and educandus and, in most cases, through peer-group interaction" (Holmberg, 1983, pp. 114). In DE, this communication can take place in real face-to-face conditions, so it is the spirit and atmosphere of conversation that should characterize educational endeavors. Discussion fora in LMSs bring together educans who study at a distance, satisfying some of the postulates of Holmberg's theory:

1. Feelings of personal relation between the teaching and learning parties promote study pleasure and motivation. Such feelings can be fostered by well-developed self-instructional material and two-way communication at a distance,
2. Intellectual pleasure and study motivation are favorable to the attainment of study goals and the use of proper study processes and methods,
3. The atmosphere, language and conventions of friendly conversation favor feelings of personal relation according to postulate 1,
4. Messages given and received in conversational forms are comparatively easily understood and remembered.

Despite the fundamental advances of the technological media used to deliver DE, these postulates remain relevant since, at a human level, the quality of interaction is a key element of effective learning. In an online learning experience, the sense of belonging, which can be reinforced via forum communication, can help students to fully and meaningfully participate in their learning procedure (DiGiacomo et al., 2023). In addition, social presence is a predictor of knowledge retention and satisfaction (Lee & Lim, 2023). Ideally, high voluntary participation in communication fora would benefit the learning community and allow tutors to closely monitor learning behavior to take targeted actions to support learners.

The students' profiles and learning style set the basis for the actions and the learning approaches they choose to follow. This study concludes that by leveraging the imprint of these actions in an LMS and using metrics of SNA, differences can be spotted in the communicational patterns that go beyond simple participation recording. The focus is on identifying patterns of student behavior through social network analysis (SNA) rather than directly correlating these behaviors with academic performance. Hopefully, the contribution of our work lies in its potential to inform future research that could establish these links more definitively. Moreover, the data collected and analyzed were not designed to measure learning outcomes directly. Although HITS and PageRank algorithms were created with completely different targeting, it is shown that they can also reveal methodological features in students' communicational approach.

This study aims to present these findings as contributions to the ongoing conversation in educational research, rather than definitive statements on the nature of forum use in distance learning. In the future, we aim to study the relationship between students' SNA metrics and students' personalities, hoping to contribute to improving the understanding of the learning process in DE.

Appendix Correlation Table

Course A-			
Variable A	Variable B	Correlation value	p value
WA1	WA2	0,385	0,156
WA1	WA3	-0,011	0,968
WA1	WA4	-0,130	0,644
WA1	In-degree	0,263	0,344
WA1	Out-degree	-0,149	0,595

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

WA1	Degree	0,235	0,400
WA1	Weighted In-degree	0,230	0,410
WA1	Weighted Out-degree	-0,040	0,887
WA1	Weighted Degree	0,179	0,523
WA1	Eccentricity	-0,730	0,002*
WA1	Closeness centrality	-0,045	0,873
WA1	Harmonic closeness centrality	-0,082	0,773
WA1	Betweenness centrality	0,071	0,800
WA1	Authority	0,218	0,436
WA1	Hub	0,106	0,706
WA1	PageRank	0,240	0,389
WA1	Eigenvector centrality	0,179	0,523
WA1	Av. WA	0,125	0,658
WA2	WA3	0,245	0,379
WA2	WA4	-0,076	0,788
WA2	In-degree	0,309	0,263
WA2	Out-degree	-0,644	0,010*
WA2	Degree	0,013	0,964
WA2	Weighted In-degree	0,231	0,406
WA2	Weighted Out-degree	-0,434	0,106
WA2	Weighted Degree	0,034	0,903
WA2	Eccentricity	-0,335	0,222
WA2	Closeness centrality	-0,393	0,148
WA2	Harmonic closeness centrality	-0,391	0,149
WA2	Betweenness centrality	0,110	0,696
WA2	Authority	0,275	0,321
WA2	Hub	-0,788	0,000*
WA2	PageRank	0,292	0,292
WA2	Eigenvector centrality	0,202	0,470
WA2	Av. WA	0,249	0,370
WA3	WA4	0,643	0,010*
WA3	In-degree	0,108	0,703
WA3	Out-degree	-0,380	0,162
WA3	Degree	-0,083	0,770
WA3	Weighted In-degree	-0,090	0,750
WA3	Weighted Out-degree	-0,583	0,023*
WA3	Weighted Degree	-0,292	0,292
WA3	Eccentricity	-0,005	0,986
WA3	Closeness centrality	-0,222	0,427

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

WA3	Harmonic closeness centrality	-0,210	0,452
WA3	Betweenness centrality	0,182	0,517
WA3	Authority	0,200	0,476
WA3	Hub	-0,170	0,544
WA3	PageRank	0,081	0,775
WA3	Eigenvector centrality	-0,043	0,880
WA3	Av. WA	0,783	0,001*
WA4	In-degree	-0,206	0,460
WA4	Out-degree	0,106	0,708
WA4	Degree	-0,191	0,495
WA4	Weighted In-degree	-0,348	0,203
WA4	Weighted Out-degree	-0,296	0,284
WA4	Weighted Degree	-0,403	0,136
WA4	Eccentricity	0,332	0,226
WA4	Closeness centrality	0,315	0,253
WA4	Harmonic closeness centrality	0,327	0,234
WA4	Betweenness centrality	0,130	0,644
WA4	Authority	-0,045	0,874
WA4	Hub	0,159	0,572
WA4	PageRank	-0,291	0,293
WA4	Eigenvector centrality	-0,389	0,152
WA4	Av. WA	0,927	0,000*
In-degree	Out-degree	-0,578	0,024*
In-degree	Degree	0,889	0,000*
In-degree	Weighted In-degree	0,900	0,000*
In-degree	Weighted Out-degree	-0,144	0,608
In-degree	Weighted Degree	0,706	0,003*
In-degree	Eccentricity	-0,652	0,008*
In-degree	Closeness centrality	-0,766	0,001*
In-degree	Harmonic closeness centrality	-0,782	0,001*
In-degree	Betweenness centrality	-0,055	0,845
In-degree	Authority	0,807	0,000*
In-degree	Hub	-0,390	0,150
In-degree	PageRank	0,946	0,000*
In-degree	Eigenvector centrality	0,576	0,025*
In-degree	Av. WA	-0,047	0,868
Out-degree	Degree	-0,140	0,618
Out-degree	Weighted In-degree	-0,296	0,284
Out-degree	Weighted Out-degree	0,801	0,000*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024

doi:10.20944/preprints202403.0638.v1

Out-degree	Weighted Degree	0,047	0,868
Out-degree	Eccentricity	0,520	0,047*
Out-degree	Closeness centrality	0,757	0,001*
Out-degree	Harmonic closeness centrality	0,764	0,001*
Out-degree	Betweenness centrality	0,149	0,595
Out-degree	Authority	-0,546	0,035*
Out-degree	Hub	0,653	0,008*
Out-degree	PageRank	-0,575	0,025*
Out-degree	Eigenvector centrality	-0,104	0,713
Out-degree	Av. WA	-0,142	0,614
Degree	Weighted In-degree	0,926	0,000*
Degree	Weighted Out-degree	0,274	0,322
Degree	Weighted Degree	0,883	0,000*
Degree	Eccentricity	-0,500	0,058
Degree	Closeness centrality	-0,504	0,055
Degree	Harmonic closeness centrality	-0,520	0,047*
Degree	Betweenness centrality	0,017	0,953
Degree	Authority	0,673	0,006*
Degree	Hub	-0,107	0,704
Degree	PageRank	0,826	0,000*
Degree	Eigenvector centrality	0,640	0,010*
Degree	Av. WA	-0,137	0,627
Weighted In-degree	Weighted Out-degree	0,242	0,385
Weighted In-degree	Weighted Degree	0,933	0,000*
Weighted In-degree	Eccentricity	-0,592	0,020*
Weighted In-degree	Closeness centrality	-0,688	0,005*
Weighted In-degree	Harmonic closeness centrality	-0,703	0,003*
Weighted In-degree	Betweenness centrality	-0,097	0,730
Weighted In-degree	Authority	0,631	0,012*
Weighted In-degree	Hub	-0,342	0,213
Weighted In-degree	PageRank	0,837	0,000*
Weighted In-degree	Eigenvector centrality	0,803	0,000*
Weighted In-degree	Av. WA	-0,226	0,419
Weighted Out-degree	Weighted Degree	0,574	0,025*
Weighted Out-degree	Eccentricity	0,180	0,522
Weighted Out-degree	Closeness centrality	0,317	0,249
Weighted Out-degree	Harmonic closeness centrality	0,317	0,250
Weighted Out-degree	Betweenness centrality	0,040	0,887

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Weighted Out-degree	Authority	-0,293	0,289
Weighted Out-degree	Hub	0,350	0,200
Weighted Out-degree	PageRank	-0,192	0,493
Weighted Out-degree	Eigenvector centrality	0,315	0,252
Weighted Out-degree	Av. WA	-0,457	0,086
Weighted Degree	Eccentricity	-0,433	0,107
Weighted Degree	Closeness centrality	-0,463	0,083
Weighted Degree	Harmonic closeness centrality	-0,476	0,073
Weighted Degree	Betweenness centrality	-0,067	0,812
Weighted Degree	Authority	0,423	0,116
Weighted Degree	Hub	-0,159	0,573
Weighted Degree	PageRank	0,635	0,011*
Weighted Degree	Eigenvector centrality	0,795	0,000*
Weighted Degree	Av. WA	-0,360	0,188
Eccentricity	Closeness centrality	0,527	0,044*
Eccentricity	Harmonic closeness centrality	0,575	0,025*
Eccentricity	Betweenness centrality	0,264	0,342
Eccentricity	Authority	-0,467	0,079
Eccentricity	Hub	0,046	0,870
Eccentricity	PageRank	-0,626	0,013*
Eccentricity	Eigenvector centrality	-0,454	0,089
Eccentricity	Av. WA	0,094	0,740
Closeness centrality	Harmonic closeness centrality	0,998	0,000*
Closeness centrality	Betweenness centrality	0,154	0,584
Closeness centrality	Authority	-0,574	0,025*
Closeness centrality	Hub	0,654	0,008*
Closeness centrality	PageRank	-0,724	0,002*
Closeness centrality	Eigenvector centrality	-0,530	0,042*
Closeness centrality	Av. WA	0,132	0,639
Harmonic closeness centrality	Betweenness centrality	0,176	0,531
Harmonic closeness centrality	Authority	-0,583	0,023*
Harmonic closeness centrality	Hub	0,627	0,012*
Harmonic closeness centrality	PageRank	-0,741	0,002*
Harmonic closeness centrality	Eigenvector centrality	-0,542	0,037*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Harmonic closeness centrality	Av. WA	0,139	0,620
Betweenness centrality	Authority	0,123	0,661
Betweenness centrality	Hub	-0,106	0,706
Betweenness centrality	PageRank	-0,117	0,678
Betweenness centrality	Eigenvector centrality	-0,059	0,835
Betweenness centrality	Av. WA	0,178	0,525
Authority	Hub	-0,323	0,240
Authority	PageRank	0,670	0,006*
Authority	Eigenvector centrality	0,383	0,159
Authority	Av. WA	0,092	0,743
Hub	PageRank	-0,357	0,191
Hub	Eigenvector centrality	-0,266	0,337
Hub	Av. WA	-0,045	0,873
PageRank	Eigenvector centrality	0,588	0,021*
PageRank	Av. WA	-0,130	0,644
Eigenvector centrality	Av. WA	-0,264	0,341

Course B-Correlation Table

Variable A	Variable B	Correlation value	p value
WA1	WA2	0,596	0,003*
WA1	WA3	0,471	0,027*
WA1	In-degree	-0,408	0,060
WA1	Out-degree	0,152	0,501
WA1	Degree	-0,347	0,114
WA1	Weighted In-degree	-0,393	0,070
WA1	Weighted Out-degree	0,163	0,469
WA1	Weighted Degree	-0,309	0,162
WA1	Eccentricity	0,296	0,182
WA1	Closeness centrality	0,207	0,356
WA1	Harmonic closeness centrality	0,218	0,329
WA1	Betweenness centrality	0,154	0,493
WA1	Authority	-0,355	0,105
WA1	Hub	0,270	0,224
WA1	PageRank	-0,448	0,037*
WA1	Eigenvector centrality	-0,513	0,015*
WA1	Av. WA	0,731	0,000*
WA2	WA3	0,718	0,000*
WA2	In-degree	-0,375	0,085
WA2	Out-degree	0,164	0,466

WA2	Degree	-0,313	0,156
WA2	Weighted In-degree	-0,345	0,116
WA2	Weighted Out-degree	0,204	0,362
WA2	Weighted Degree	-0,254	0,253
WA2	Eccentricity	0,329	0,135
WA2	Closeness centrality	0,258	0,247
WA2	Harmonic closeness centrality	0,269	0,226
WA2	Betweenness centrality	0,151	0,503
WA2	Authority	-0,225	0,314
WA2	Hub	0,330	0,133
WA2	PageRank	-0,433	0,044*
WA2	Eigenvector centrality	-0,432	0,045*
WA2	Av. WA	0,914	0,000*
WA3	In-degree	-0,133	0,556
WA3	Out-degree	0,156	0,487
WA3	Degree	-0,084	0,711
WA3	Weighted In-degree	-0,069	0,759
WA3	Weighted Out-degree	0,202	0,368
WA3	Weighted Degree	-0,008	0,971
WA3	Eccentricity	0,194	0,386
WA3	Closeness centrality	0,215	0,336
WA3	Harmonic closeness centrality	0,217	0,332
WA3	Betweenness centrality	0,180	0,424
WA3	Authority	0,029	0,899
WA3	Hub	0,291	0,189
WA3	PageRank	-0,153	0,496
WA3	Eigenvector centrality	-0,116	0,607
WA3	Av. WA	0,900	0,000*
In-degree	Out-degree	0,037	0,870
In-degree	Degree	0,962	0,000*
In-degree	Weighted In-degree	0,964	0,000*
In-degree	Weighted Out-degree	0,121	0,591
In-degree	Weighted Degree	0,896	0,000*
In-degree	Eccentricity	-0,463	0,030*
In-degree	Closeness centrality	-0,563	0,006*
In-degree	Harmonic closeness centrality	-0,568	0,006*
In-degree	Betweenness centrality	0,188	0,402
In-degree	Authority	0,958	0,000*
In-degree	Hub	-0,202	0,367

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

In-degree	PageRank	0,963	0,000*
In-degree	Eigenvector centrality	0,855	0,000*
In-degree	Av. WA	-0,325	0,140
Out-degree	Degree	0,307	0,165
Out-degree	Weighted In-degree	0,122	0,588
Out-degree	Weighted Out-degree	0,883	0,000*
Out-degree	Weighted Degree	0,345	0,115
Out-degree	Eccentricity	0,567	0,006*
Out-degree	Closeness centrality	0,394	0,069
Out-degree	Harmonic closeness centrality	0,414	0,055
Out-degree	Betweenness centrality	0,551	0,008*
Out-degree	Authority	0,064	0,777
Out-degree	Hub	0,871	0,000*
Out-degree	PageRank	0,011	0,961
Out-degree	Eigenvector centrality	0,004	0,987
Out-degree	Av. WA	0,182	0,417
Degree	Weighted In-degree	0,951	0,000*
Degree	Weighted Out-degree	0,355	0,105
Degree	Weighted Degree	0,947	0,000*
Degree	Eccentricity	-0,287	0,196
Degree	Closeness centrality	-0,429	0,047*
Degree	Harmonic closeness centrality	-0,429	0,046*
Degree	Betweenness centrality	0,329	0,135
Degree	Authority	0,929	0,000*
Degree	Hub	0,044	0,844
Degree	PageRank	0,920	0,000*
Degree	Eigenvector centrality	0,816	0,000*
Degree	Av. WA	-0,260	0,243
Weighted In-degree	Weighted Out-degree	0,263	0,238
Weighted In-degree	Weighted Degree	0,966	0,000*
Weighted In-degree	Eccentricity	-0,369	0,091
Weighted In-degree	Closeness centrality	-0,432	0,045*
Weighted In-degree	Harmonic closeness centrality	-0,438	0,042*
Weighted In-degree	Betweenness centrality	0,181	0,420
Weighted In-degree	Authority	0,935	0,000*
Weighted In-degree	Hub	-0,113	0,615
Weighted In-degree	PageRank	0,928	0,000*
Weighted In-degree	Eigenvector centrality	0,909	0,000*
Weighted In-degree	Av. WA	-0,278	0,210

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Weighted Out-degree	Weighted Degree	0,503	0,017*
Weighted Out-degree	Eccentricity	0,478	0,024*
Weighted Out-degree	Closeness centrality	0,350	0,110
Weighted Out-degree	Harmonic closeness centrality	0,367	0,093
Weighted Out-degree	Betweenness centrality	0,365	0,095
Weighted Out-degree	Authority	0,140	0,533
Weighted Out-degree	Hub	0,727	0,000*
Weighted Out-degree	PageRank	0,027	0,904
Weighted Out-degree	Eigenvector centrality	0,088	0,697
Weighted Out-degree	Av. WA	0,224	0,317
Weighted Degree	Eccentricity	-0,203	0,365
Weighted Degree	Closeness centrality	-0,293	0,185
Weighted Degree	Harmonic closeness centrality	-0,294	0,184
Weighted Degree	Betweenness centrality	0,260	0,243
Weighted Degree	Authority	0,875	0,000*
Weighted Degree	Hub	0,093	0,681
Weighted Degree	PageRank	0,839	0,000*
Weighted Degree	Eigenvector centrality	0,838	0,000*
Weighted Degree	Av. WA	-0,189	0,399
Eccentricity	Closeness centrality	0,720	0,000*
Eccentricity	Harmonic closeness centrality	0,759	0,000*
Eccentricity	Betweenness centrality	0,411	0,058
Eccentricity	Authority	-0,365	0,095
Eccentricity	Hub	0,708	0,000*
Eccentricity	PageRank	-0,400	0,065
Eccentricity	Eigenvector centrality	-0,379	0,082
Eccentricity	Av. WA	0,306	0,166
Closeness centrality	Harmonic closeness centrality	0,998	0,000*
Closeness centrality	Betweenness centrality	0,032	0,889
Closeness centrality	Authority	-0,492	0,020*
Closeness centrality	Hub	0,566	0,006*
Closeness centrality	PageRank	-0,517	0,014*
Closeness centrality	Eigenvector centrality	-0,388	0,074
Closeness centrality	Av. WA	0,264	0,235
Harmonic closeness centrality	Betweenness centrality	0,057	0,800
Harmonic closeness centrality	Authority	-0,495	0,019*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Harmonic closeness centrality	Hub	0,588	0,004*
Harmonic closeness centrality	PageRank	-0,520	0,013*
Harmonic closeness centrality	Eigenvector centrality	-0,396	0,068
Harmonic closeness centrality	Av. WA	0,272	0,221
Betweenness centrality	Authority	0,289	0,191
Betweenness centrality	Hub	0,542	0,009*
Betweenness centrality	PageRank	0,190	0,396
Betweenness centrality	Eigenvector centrality	-0,024	0,914
Betweenness centrality	Av. WA	0,189	0,401
Authority	Hub	-0,125	0,580
Authority	PageRank	0,920	0,000*
Authority	Eigenvector centrality	0,833	0,000*
Authority	Av. WA	-0,171	0,447
Hub	PageRank	-0,180	0,424
Hub	Eigenvector centrality	-0,210	0,349
Hub	Av. WA	0,347	0,114
PageRank	Eigenvector centrality	0,897	0,000*
PageRank	Av. WA	-0,369	0,091
Eigenvector centrality	Av. WA	-0,367	0,093

Course A-SNA normalized metrics

Label	In-degree	Out-degree	Degree	Weighted In-degree	Weighted Out-degree	Weighted Degree	Eccentricity	Closness centrality	Harmonic closness centrality	Betweenness centrality	Authority	Hub	PageRank	Eigenvector centrality
Ast14	0	1	0,333	0,000	0,667	0,125	0,250	1,000	1,000	0,000	0,000	1,000	0,000	0,000
Ast11	0	0,5	0,000	0,000	0,333	0,000	0,500	0,571	0,625	0,000	0,000	0,000	0,000	0,000

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024
doi:10.20944/preprints202403.0638.v1

Ast12	Ast13	Ast5	Ast6	Ast7	Ast16	Ast8	Ast1	Ast9	Ast2	Ast3	Ast10
0,5	0,75	0,25	0	0,5	0,25	0	1	0	0	0,5	0,75
0	0,5	0	0,5	0	0	0,5	0	0,5	0,5	0,5	0
0,333	1,000	0,000	0,000	0,333	0,000	0,000	1,000	0,000	0,000	0,667	0,667
0,333	1,000	0,167	0,000	0,333	0,167	0,000	0,833	0,000	0,000	0,833	0,500
0,000	1,000	0,000	0,333	0,000	0,000	0,333	0,000	0,333	0,333	0,667	0,000
0,125	1,000	0,000	0,000	0,125	0,000	0,000	0,500	0,000	0,000	0,750	0,250
0,000	0,000	0,000	0,250	0,000	0,000	0,250	0,000	0,500	1,000	0,000	0,000
0,000	0,000	0,000	1,000	0,000	0,000	1,000	0,000	0,571	0,409	0,000	0,000
0,000	0,000	0,000	1,000	0,000	0,000	1,000	0,000	0,625	0,481	0,000	0,000
0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
0,689	0,452	0,163	0,000	0,689	0,034	0,000	1,000	0,000	0,000	0,163	0,208
0,000	0,000	0,000	0,374	0,000	0,000	0,172	0,000	0,000	0,000	0,000	0,000
0,299	0,605	0,471	0,000	0,299	0,153	0,000	1,000	0,000	0,000	0,471	0,841
0,134	0,509	0,390	0,000	0,134	0,006	0,000	0,530	0,000	0,000	1,000	0,019

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Ast4
0,25
0,5
0,333
0,167
0,333
0,125
0,500
0,571
0,625
1,000
0,396
0,000
0,146
0,128

Course B-SNA normalized metrics

Label	Bst14	Bst9	Bst18	Bst13	Bst8	Bst22	Bst12	Bst17
In-degree	0,111	0,667	0,000	0,000	0,333	0,000	0,444	0,000
Out-degree	0,333	0,333	0,333	0,333	0,667	1,000	1,000	0,333
Degree	0,111	0,667	0,000	0,000	0,444	0,222	0,667	0,000
Weighted In-degree	0,077	0,462	0,000	0,000	0,538	0,000	0,385	0,000
Weighted Out-degree	0,250	0,250	0,250	0,250	1,000	0,750	0,750	0,250
Weighted Degree	0,077	0,462	0,000	0,000	0,769	0,154	0,538	0,000
Eccentricity	0,500	0,000	0,500	0,500	0,500	0,500	1,000	0,500
Closness centrality	1,000	0,000	1,000	1,000	1,000	1,000	0,600	1,000
Harmonic closness centrality	1,000	0,000	1,000	1,000	1,000	1,000	0,667	1,000
Betweenness centrality	0,026	0,000	0,000	0,000	0,026	0,000	1,000	0,000
Authority	0,245	0,643	0,000	0,000	0,311	0,000	0,572	0,000
Hub	0,216	0,000	0,495	0,180	0,495	1,000	0,966	0,180
PageRank	0,022	0,441	0,000	0,000	0,120	0,000	0,365	0,000
clustering	0,000	0,200	0,000	0,000	0,333	0,167	0,167	0,000
Eigenvector centrality	0,005	0,236	0,000	0,000	0,236	0,000	0,072	0,000

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Bst15	Bst3	Bst16	Bst19	Bst10	Bst11	Bst4	Bst5	Bst6	Bst20	Bst7	Bst21
0,000	0,667	0,000	0,444	0,556	0,000	0,111	0,000	0,111	1,000	0,000	0,444
0,333	0,667	0,333	0,333	0,333	0,333	0,000	0,333	0,333	0,333	0,667	0,000
0,000	0,778	0,000	0,444	0,556	0,000	0,000	0,000	0,111	1,000	0,111	0,333
0,000	0,615	0,000	0,385	0,385	0,000	0,077	0,000	0,077	1,000	0,000	0,308
0,250	0,750	0,250	0,500	0,250	0,250	0,000	0,250	0,250	0,250	0,750	0,000
0,000	0,769	0,000	0,462	0,385	0,000	0,000	0,000	0,077	1,000	0,154	0,231
0,500	0,500	0,500	0,000	0,000	0,500	0,000	0,500	0,000	0,000	1,000	0,000
1,000	1,000	1,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000	0,667	0,000
1,000	1,000	1,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000	0,750	0,000
0,000	0,079	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
0,000	0,584	0,000	0,534	0,435	0,000	0,256	0,000	0,000	1,000	0,000	0,363
0,000	0,495	0,216	0,000	0,000	0,319	0,000	0,319	0,000	0,000	0,528	0,000
0,000	0,494	0,000	0,214	0,418	0,000	0,042	0,000	0,000	1,000	0,000	0,419
0,000	0,238	0,000	0,333	0,200	0,000	0,000	0,000	0,000	0,222	0,000	0,000
0,000	0,329	0,000	0,189	0,231	0,000	0,056	0,000	0,024	1,000	0,000	0,072

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

Bst1	Bst2
0,222	0,000
0,333	0,333
0,222	0,000
0,154	0,000
0,250	0,250
0,154	0,000
0,000	1,000
0,000	0,667
0,000	0,750
0,000	0,000
0,000	0,000
0,000	0,289
0,151	0,000
0,500	0,000
0,047	0,000

References

1. Adraoui, M., Retbi, A., Idrissi, M. K., & Bennani, S. (2017). Social learning analytics to describe the learners' interaction in online discussion forum in Moodle. In 2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET) (pp. 1-6). IEEE.
2. Borgatti, S. P., & Halgin, D. S. (2011). On network theory. Organization science, 22(5), 1168-1181.
3. Brin, S., & Page, L. (1998, Apr). The Anatomy of a Large-Scale Hypertextual Web Search Engine. Paper presented at the Proceedings of the seventh international conference on World Wide Web 7, Brisbane, Australia, from http://dbpubs.stanford.edu/pub/1998-8
4. Brindley, J.E., Walti, C., & Blaschke, L. M. (2009). Creating Effective Collaborative Learning Groups in an Online Environment. IRRODL, 10(3). Retrieved 10-01-2014 from http://www.irrodl.org/index.php/irrodl/article/view/675/1271
5. Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS. IEEE Transactions on Learning Technologies, 10(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312
6. Da Silva, L. F. C., Barbosa, M. W., & Gomes, R. R. (2019). Measuring participation in distance education online discussion forums using social network analysis. Journal of the Association for Information Science and Technology, 70(2), 140-150.
7. Davis, B., Carmean, C., & Wagner, E. (2009). "The Evolution of the LMS:From Management to Learning-Deep Analysis of Trends Shaping the Future of e-Learning". Santa Rosa, CA: The ELearning Guild Research.
8. Davis. Eigenfactor (2008) "Does the Principle of Repeated Improvement Result in Better Journal Impact Estimates than Raw Citation Counts? Journal of the American Society for Information Science and Technology 59(13): 2186-2188. DOI:http://dx.doi.org/10.1002/asi.20943
9. DiGiacomo, D. K., Usher, E. L., Han, J., Abney, J. M., Cole, A. E., & Patterson, J. T. (2023). The benefits of belonging: Students' perceptions of their online learning experiences. Distance Education, 1-16.
10. Easley & J. Kleingeld. (2010). Networks, Crowds, and Markets: Reasoning about a Highly Connected World. Cambridge University Press. file:///C:/Users/user/Downloads/networks-book.pdf
11. Geigle, & C. Zhai, (2017). "Modeling MOOC Student Behavior with Two-Layer Hidden Markov Models". In the Proceedings of the 4th ACM Conference on Learning at Scale. April 2017 pp. 205–208. DOI: https://doi.org/10.1145/3051457.3053986
12. Gkontzis, A.F., Kotsiantis, S., Kalles, D., Panagiotakopoulos, C.T., & Verykios, V.S. (2020). Polarity, emotions and online activity of students and tutors as features in predicting grades. Intelligent Decision Technologies, 14(3), pp. 409-436. DOI: 10.3233/IDT-190137
13. Hernández García, Á., González González, I., Jiménez Zarco, A. I., & Chaparro Peláez, J. (2016). Visualizations of online course interactions for social network learning analytics.
14. Hernández-García, Á., González-González, I., Jiménez-Zarco, A. I., & Chaparro-Peláez, J. (2015). Applying social learning analytics to message boards in online distance learning: A case study. Computers in Human Behavior, 47, 68-80.
15. Holmberg, B. (1983). Guided didactic conversation in distance education. In D. Sewart, D. Keegan, and B. Holmberg (Eds.), Distance education: International perspectives (pp. 114-122). London: Croom Helm.
16. Huang, J., Dasgupta, A., Ghosh, A., Manning, J., & Sanders, M. (2014, March). Superposter behavior in MOOC fora. In Proceedings of the first ACM conference on Learning@ scale conference (pp. 117-126).
17. Kagklis, A. Karatrantou, C. T. Panagiotakopoulos & V. S. Verykios (2015). A learning analytics methodology for detecting sentiment in student fora: a case study in distance education. European Journal of Open, Distance and e-Learning 18(2). DOI: 10.1515/eurodl-2015-0014
18. Kagklis, V., Karatrantou, A., Tantoula, M., Panagiotakopoulos, C., Verykios, V.S. (2015). A Learning Analytics Methodology for Detecting Sentiment in Student Fora: A Case Study in Distance Education. European Journal of Open, Distance and E-learning, 18(2), pp. 74-94.Writer's Handbook. Mill Valley, CA: University Science, 1989.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

19. Kipling, R. P., Stiles, W. A., de Andrade-Lima, M., MacKintosh, N., Roberts, M. W., Williams, C. L., ... & Watson-Jones, S. J. (2023). Interaction in online postgraduate learning: what makes a good forum?. Distance Education, 44(1), 162-189.
20. Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., & Tomkins, A. S. (1999). The web as a graph: Measurements, models and methods. In International Computing and Combinatorics Conference (pp. 1-17). Springer, Berlin, Heidelberg.
21. Kleinberg, Jon (1999). "Authoritative sources in a hyperlinked environment" (PDF). Journal of the ACM. 46 (5): 604–632. Retreived from http://www.cs.cornell.edu/home/kleinber/auth.pdf
22. Lang, C., Siemens, G., Wise, A., & Gasevic, D. (2017). Handbook of Learning Analytics. Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17
23. Lee, M., & Lim, J. (2023). Do online teaching and social presences contribute to motivational growth?. Distance Education, 44(1), 66-85.
24. Lee, S. Y., Chae, H. S., & Natriello, G. (2018). Identifying User Engagement Patterns in an Online Video Discussion Platform. International Educational Data Mining Society.
25. Livingston, J. A. (2003). Metacognition: An Overview. Psychology, 13, 259-266.
26. López-Flores, N. G., Óskarsdóttir, M., & Islind, A. S. (2022, March). Analysis of discussion forum interactions for different teaching modalities based on temporal social networks. In Proceedings of the NetSciLA22 workshop.
27. Miyazoe, T., & Anderson, T. (2013). Interaction equivalency in an OER, MOOCS and informal learning era.
28. Moore, M.G. (1973). Toward a Theory of Independent Learning and Teaching. Journal of Higher Education, 44, 661-679.
29. Motz, J. D. Quick, J. A. Wernert & T. A. Miles (2021). "A pandemic of busywork: Increased online coursework following the transition to remote instruction is associated with reduced academic achievement". Online Learning, 25(1), 70-85.https://doi.org/10.24059/olj.v25i1.2475
30. Panagiotakopoulos, T. Tsiatsos, A. Lionarakis, N. Tsanakos (2013). Teleconference in support of distance learning: Views of educators. Open Education-The Journal for Open and Distance Education and Educational Technology, 9(1), pp. 5-18. Open Education, ISSN: 1791-9312.
31. Perraton, H. (1988). A theory for distance education. In D. Sewart, D. Keegan & B. Holmberg (eds), Distance Education: International perspectives, (pp. 34-45). New York: Routledge.
32. Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. Information Processing and Management, 12(5), 297-312.
33. Putnik, G., Costa, E., Alves, C., Castro, H., Varela, L., & Shah, V. (2016). Analysing the correlation between social network analysis measures and performance of students in social network-based engineering education. International Journal of Technology and Design Education, 26, 413-437.
34. Rigney, D. (2010). *The Matthew effect: How advantage begets further advantage*. Columbia University Press.
35. Sanchez, T., Naranjo, D., Vidal, J., Salazar, D., Pérez, C., & Jaramillo, M. (2021). Analysis of Academic Performance Based on Sociograms: A Case Study with Students from At-Risk Groups. Journal of Technology and Science Education, 11(1), 167-179.
36. Saqr, M., Elmoazen, R., Tedre, M., López-Pernas, S., & Hirsto, L. (2022). How well centrality measures capture student achievement in computer-supported collaborative learning?–A systematic review and meta-analysis. Educational Research Review, 100437.
37. Saqr, M., Elmoazen, R., Tedre, M., López-Pernas, S., & Hirsto, L. Educational Research Review.
38. 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.
39. Shonyy, (2021, January 8). PageRank: Link Analysis Explanation and Python Implementation from Scratch-The Algorithm that Starts Google. Towards Data Science. Retreived from https://towardsdatascience.com/pagerank-3c568a7d2332
40. Siemens, G., & Baker, R.S.J. (2012). Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. LAK12, 2012.
41. Simpson, O. (2002). Supporting Students in open and Distance Learning. London: KoganPage.
42. Sturludóttir, E. G., Arnardóttir, E., Hjálmtýsson, G., & Óskarsdóttir, M. (2021). Gaining insights on student course selection in higher education with community detection. arXiv preprint arXiv:2105.01589.
43. Tsoni, R., Panagiotakopoulos, T.C., & Verykios, V.S. (2021). Revealing latent traits in the social behavior of distance learning students. Education and Information Technologies [Online]. https://link.springer.com/article/10.1007%2Fs10639-021-10742-6
44. Tsoni, R., Paxinou, E., Stavropoulos, E., Panagiotakopoulos, C. T., & Verykios V. S. (2020). Looking under the hood of students' collaboration networks in distance learning. The Envisioning Report for Empowering Universities, 39

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2024 doi:10.20944/preprints202403.0638.v1

45. Tsoni, R., Sakkopoulos, E., & Verykios, V. S. (2022). Revealing latent student traits in distance learning through SNA and PCA. In Handbook on Intelligent Techniques in the Educational Process: Vol 1 Recent Advances and Case Studies (pp. 185-209). Cham: Springer International Publishing.
46. Tsoni, R., Sakkopoulos, E., Panagiotakopoulos, C.T., & Verykios, V.S. (2021). On the equivalence between bimodal and unimodal students' collaboration networks in distance learning. Intell. Decis. Technol. 15(2), pp. 305-319. DOI: 10.3233/IDT-200137
47. Tsoni, R., Samaras, C., Paxinou, E., Panagiotakopoulos, C., & Verykios, V.S. (2019). From analytics to cognition: Expanding the reach of data in learning. In the Proceedings of the 11th International Conference on Computer Supported Education, 2, pp. 458-465. DOI: 10.5220/0007751904580465
48. Vygotsky, L. (1987). Zone of proximal development. Mind in society: The development of higher psychological processes, 5291, 157.
49. Wedemeyer, C.A. (1977). Independent study. In A.S. Knowles (ed.), The International Encyclopedia of Higher Education. Boston: CIHED.
50. Wollny, S., Di Mitri, D., Jivet, I., Muñoz-Merino, P., Scheffel, M., Schneider, J., ... & Drachsler, H. (2023). Students' expectations of Learning Analytics across Europe. Journal of computer assisted learning.
51. Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. Educational psychologist, 25(1), 3-17.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.