

RESEARCH

Open Access



Uncovering trend-based research insights on teaching and learning in big data

Young-Eun Park* 

*Correspondence:
ypark@psu.edu.sa
Management Department,
College of Business
Administration, Prince Sultan
University, Riyadh, Saudi
Arabia

Abstract

Along with the big data era, digital transformation has had a transformative effect on modern education tremendously in higher education. It transforms an institutional core value of education to better meet students' needs by leveraging big data and digital technology. Based on this background, this study attempts to catch the principal trends, or new directions, paradigms as predictors with an association of each topic by discovering the up-to-date research trends on teaching and learning in higher education via text mining techniques. For this, 285 research articles in the area of teaching and learning in higher education were collected from several big databases (distinguishable publishers' web platforms) through search engines for 2 years in 2018–2019. Then it was analyzed using a semantic network analysis that processes natural human language. Consequently, research results show a relatively high connection with 'student' or 'student-centered/led' rather than 'teacher-led'. Moreover, it exhibits that the practice and assessment in learning can be attained via diverse learning activities, containing community or outreach activities. Besides, research in academic contexts, experience-based classes, the effect of group activities, how students' feelings or perceptions, and relationships affect learning outcomes were addressed as the main topics through topic modeling of LDA, a machine learning algorithm. This study proposes that educators, researchers, and even academic leaders can exert extraordinary power to reshape educational quality programs for future education and in a timely manner with recognizable trends or agendas in teaching and learning of higher education.

Keywords: Teaching and learning, Research trends, Higher education, Big data, Semantic network analysis

Introduction

Digital transformation in higher education covers many things from using digital tools such as LMS (Learning Management System), Interactive whiteboard, various application tools for e-learning, etc. in university classrooms to digitize university documents. However, it does not just involve in utilizing advanced tools. The change and growing demand can be more profound and deep, incorporating whole aspects of education. Although there is an initial resistance to new technologies caused by continually changing, we have no choice but to hold and follow this latest trend [11, 53]. Moreover, no matter what has happened to us in the past and what is going on in our current lives, there is no power to keep us from having an unknown future with high uncertainty. It

is the main reason we should be aware of the latest trends or changing circumstances around us. Accordingly, efforts to look at trends or predict a new paradigm based on historical big data in teaching and learning has been around continuously in the various educational sectors of higher education [1, 4, 16, 17, 39, 53]. In particular, main issues, which have traditionally started from finding trends about students' learning [17, 20], expand into teaching and learning broadly [22, 58]. Furthermore, now it amplifies into research and development [7, 16, 18, 19, 27, 29, 37, 41, 42, 44]. Besides, previous studies favored measuring the higher education curriculum's achievement or improvement and measuring the outcomes to fulfill the highest standards [2, 6]. Past research trends that studied diverse factors affecting students' learning attitudes and learning outcomes have been extended to diversified research areas. Those comprehensively connect distinct educational issues and teaching and learning methods. For example, new techniques such as flipped learning, blended learning, and interactive ways maximize education effectiveness [45, 48, 49, 51, 64]. Furthermore, there is an ever-increasing range of connections with multiple stakeholders around education, management, administration, policy, and leadership in higher education [8, 35, 44].

Thus, recognizing the latest issues or trends of how teaching and learning in higher education have conducted in research has a special meaning to reflect the present and previous studies and seek desirable directions in the development of future education [7, 18, 19, 27, 29, 37, 41, 42]. Consequently, a number of researchers in diverse educational sectors have shown many distinguished studies related to those topics with different perspectives to keep up with the most popular and the latest educational trends. However, most studies are independent studies based on a specific situation or context rather than grasping the overall academic flow. Besides, most of them were investigated with qualitative and subjective methods through surveys and interviews, such as focus group interviews. Also, reviews of scientific publications such as articles or books, course syllabus, etc. are implemented to manually gather non-numerical data. This qualitative research approach provides obvious implications and contributions comprehensively to understand and explore an educational situation or a higher education phenomenon. However, it is a study confined to individual conditions or specific circumstances independently. Moreover, there is a potential to create subjective judgments or arguments when applying for different occasions or environments [41, 42]. Hence, this study aims to supplement the limitations of those existing studies and grasp the general current flow of teaching and learning through big data. By identifying the latest teaching and learning trends in higher education through semantic network analysis, a text mining technique using big data of unstructured texts, this study catches the fundamental directions and main topics, including the linkage of each issue. In conclusion, this research can provide a wealth of insights to guide educators, researchers, and academic leaders in higher education with a trend-based approach in quality education. The latest educational issues identified using big data and advanced technologies will help renovate and reshape future educational programs. Also, in the face of rapidly changing educational trends and dynamic environments, it will meet the different needs of diverse stakeholders around higher education.

The rest of this research paper is organized and illustrated as follows. The next section presents reviewing existing studies extensively. Then, a proposed research framework

with research questions is addressed. After that, the data and methodology are presented and analyzed in the section of results and discussion. Lastly, the implication and conclusion are included in a future research direction in the last section.

Literature review

Researches about teaching and learning in higher education

A past or present, a critical issue to higher education lies in teaching and learning. Accordingly, different agendas from a wide variety of perspectives in higher education are progressively evolving. However, we need to revisit the basic principals and focus on the essential subjects of teaching and learning [9, 45, 48, 49, 51, 60, 64]. Traditionally, education has focused on finding the diverse matters surrounding the roles of students and teachers, respectively, to examine the learning effects and the factors that influence it, that is, to find cause and effect (causality) in teaching and learning [16, 17]. It has progressed into an in-depth discussion of how psychological factors, such as students' perceptions, feelings, or relationships with teachers, influence learning outcomes [24, 36, 62]. What is more, many studies examined how various demographical factors such as gender, race, and income level of families affect students' learning or its outcomes [25, 30, 57]. In addition to this approach, attention was paid to the teachers' point of view to find out more effective teaching methods and what new teaching methods were being developed and used and how those methods worked. For example, flipped learning, blended learning, online learning, or interactive learning using various technology tools or simulation game. These are the most recently adopted teaching methods [19, 21, 23, 34, 60].

Other endeavors have been made to find general trends in teaching and learning with a holistic perspective [13, 16, 19, 20, 22, 27, 37, 41, 42]. Deng et al. [13] focused on identifying trends and categorizing the study on Massive Open Online Courses in teaching and learning. Elton and Laurillard [16] sought to find research trends in students' learning and discover new research paradigms. They analyzed the trends to uncover the determinants of how humans learn, the differences among individuals in human education, how content elements affect learning, and how contextual factors affect learning. Guri-Rozenblit [20] reviewed and analyzed four books that can use to examine trends in learning. Based on this, he defined the definitions of distance education and open education. He covered a wide range of free public/distance systems, course design, advanced technology, and delivery systems, student support and survival issues, and lastly, inter-university and inter-institutional collaboration issues. Henderson et al. [22] studied students' perceptions of useful digital technologies in teaching and learning in the university, which has an online education or interactive education through an online system. It is attracting attention as research that captures the transforming the nature of university education. Nikitina and Lapina [41] proposed that recent business education trends were organized into three categories: a curriculum that meets the desire of society and business, partnership & networking, and a modern and flexible teaching method in their research. Besides, new forms of teaching and learning, including blended learning, interactive learning, and flipped learning, have been addressed by many scholars [21, 34, 60]. Besides that, a large number of studies have mainly concentrated on the numerous factors or trends affecting educational development and management [19, 27,

37, 41, 42]. For instance, Nomuoja [42] studied the current trends emerging in business schools of higher education. Consequently, career awareness, risk management, people-oriented strategy and management, and skills-based curriculum were mainly discussed. Moreover, there are interviews from international top MBA schools to discover major MBA trends such as 'growing trend of double degrees,' growth acceleration of online or technology-based education and blended learning in business education [19, 23, 37]. Accordingly, a considerable amount of research has been done with broad and varied teaching and learning perspectives. However, most of them were independent studies investigated based on a specific situation or context rather than grasping the educational flow or trends. Moreover, there is still a lack of study that looks at the global direction of such research more objectively and quantitatively using big data. Hence, this study aims to fill in the gap of these existing studies.

Semantic network analysis using big data of the unstructured text

We live in an era where all aspects of our lives are uncertain and rapidly changing [28, 33, 47]. The best way to cope with this uncertain and unknown future is to predict and prepare for the future based on various historical big data by reducing this prediction error. In this regard, people focus on using big data to read trends and prepare for the unknown future. This substantial phenomenon is well represented in diverse and separate research fields as well. Many scholars in a very different area are working actively to discover insights into big data using various data mining techniques [1, 15, 26, 38, 46, 47, 52, 53, 55, 63]. We can deal with big data or data sets due to the breakthrough technology, which is too complicated or broad to be dealt with by traditional data-processing approaches. In particular, it became possible to analyze a large amount of unstructured text data through text mining, one of the data mining techniques, as linguistic techniques have developed and applied to diverse areas [11, 38, 61].

A morphological or semantic network analysis deals with dividing a sentence into the smallest meaningful unit of language, namely, morphemes by importing unstructured text data such as speeches, comments, or posting in social media like Twitter, Instagram, or any bibliographic information (for example, books, scholarly articles, records, interviews, etc.) [15, 26, 43, 47, 63]. It automatically extracts words in sentences, paragraphs, and documents to make it simple to construct a word-to-word network according to the degree of nearness or adjacency between words [3, 26]. Based on that, network structures provide intuitive and beneficial illustrations for modeling semantic inference and knowledge [55]. Through this, we can comprehend the relationship among words or understand their association by combining topics through proper interpretations in a given text [12, 50, 55]. The more commended, the larger the size of the morpheme or word. It can then be seen at first sight, as it were, to visually stress major issues or agendas such as keywords in unstructured documents to extract critical attributes, mainly in big data that manages a large amount of information. Nodes in a semantic network mean words, and links are word-to-word adjacency relationships [43]. Until recently, network analysis required data structured by nodes and ties ahead of time. The subsequent processes were performed by individual programs, which required plenty of human efforts and time. However, with state-of-the-art technology development, natural language processing is built into

data mining programs, which can directly enter unstructured text data and extract words (nodes) in morphological units and create network data encompassing words. It broadens the network analysis horizon with massive unstructured text data [12, 29]. Accordingly, a large number of scholars has ripened into a semantic network analysis as a powerful tool of text mining in numerous ways since Rice and Danowski [50] built a basic framework for network analysis [15, 40, 50, 56].

The purpose of analyzing text using text mining is very diverse. It is possible to comprehend between the lines in which the document intends to deliver by reassembling the text. Also, by visually grasping the relationship between the main concepts and other keywords in the text, it is easy to understand various networks. Through this, it is achievable to analyze the roles of words and their relationships by recognizing the word associations. One of the most significant benefits in text mining is to analyze the terms both qualitatively and quantitatively [11, 38, 39]. Additionally, it uses to visualize or illustrate the relationship between objects or people in text and topic modeling [26, 43]. For this, a large amount of information can efficiently and effectively utilize to generate more comprehensive and extended knowledge, analytical reasoning, and even explorative analysis, which is the final goal of text analysis [11, 12, 14, 38, 39, 59]. With those benefits of this approach, many scholars have discussed various topics with different perspectives using big data. Many scholars and observers have found huge opportunities and tremendous potentials of semantic network analysis with recognizing centrality indicators between words and sub-network structures of words [32, 50]. Many of those studies exhibit the possibility of the ongoing development of the semantic networks as a powerful research tool emerging with the big data era's advent. In particular, semantic network analysis is used in research to study teaching and learning in higher education. Shen and Ho [51] showed the importance of technology-enhanced learning (TEL) through a semantic approach as an inspired way to improve teaching and learning outcomes in high education. Kim [27] determined the study trends of music education using the semantic network analysis. Lee [31] analyzed the research trends in the area of journalism, pursuing the key words of the abstract of research articles published in 2005–2015 through semantic network analysis, then, finally established knowledge system as a result. Besides, Kim et al. [29] applied semantic network analysis to draw a significant agenda of nursing care service opinions by extracting data from online news and social media data. Recently, Park [47] took the data of news media and social media to compare the trends from the two different kinds of big data sources to predict leading Korean companies' sustainability.

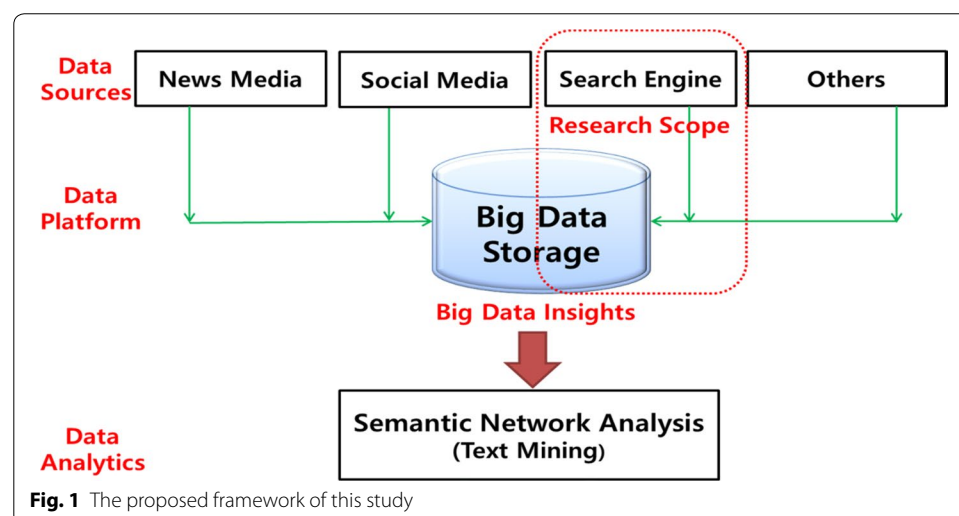
Based on those previous studies, this study aims to investigate the most recent research issues and latest trends of teaching and learning in higher education through semantic network analysis. Using a large amount of unstructured text data, we can effectively and efficiently discover trendy insights and directions of future education in teaching and learning and research [15, 52, 53, 55]. Accordingly, it expects to generate subsequent development of knowledge and intuition to comprehend a new paradigm of future education in general, which is just around the corner. It would be very constructive and beneficial to educators, researchers, academic leaders, and even administrative leaders in higher education.

Proposed research framework

To pinpoint major agendas and trends in teaching and learning of higher education, semantic network analysis, a data mining technique, was used in this study. Accordingly, there is no theoretical framework with hypotheses as the data-driven approach is used in this paper. This data-driven methodology became an extraordinarily capable and promising area. A massive amount of information reserved in electronic and digital records on the internet brings tremendous opportunities and impacts remarkably for knowledge discovery, information extraction, and analytical reasoning in education fields [15, 40, 61]. Thus, this empowers one to extract important algorithmic properties that give powerful intuitions and insights into the structure of networks and graphs [38, 53, 55, 65]. As previous literature shows, a researcher can collect big data from various sources and platforms. For example, news channels (such as BBC, CNN, ABC, etc.), social media (such as Facebook, Twitter, Instagram, YouTube, etc.), web or internet search engines (such as Google, Bing, Yahoo, AOL, Journal databases or publishers' databases, etc.), other financial reports (such as financial statements, press releases, conference calls regarding earnings and related information, etc.), and so on [11, 39, 47, 53]. In this study, the data were gathered for analysis from several distinguishable publishers (Sage Publications, Taylor & Francis, and Elsevier BV)' web platforms through search engines. Figure 1 shows the proposed framework of this study with a holistic approach.

This study attempts to determine the most recent research agendas or trends of the leading higher education journals about teaching and learning in 2018 and 2019 through semantic network analysis. As the global trend is changing very fast, this study emphasizes teaching and learning in the last 2 years. For this purpose, this study establishes the following research questions.

- (1) What are the main trends or agendas of teaching and learning in higher education in the last two years?
- (2) What are the critical attributes of teaching and learning in higher education, and what are the implications of this?
- (3) How are the specific sub-domains (topic modeling) of teaching and learning in higher education categorized as future education strategies?



Data and methodology

Data collection

This study aims to identify the most recent educational trends and predict future directions or shifts by recognizing the main issues of teaching and learning in Higher education. For that, the data collected from 285 research articles of four international ‘top-tier’ journals ranked as ISI/SCOPUS Q1 in this field for 2 years (2018–2019) in the big databases of each publisher’s web platform. The data were selected according to the following criteria using the search engine: ISI/SCOPUS indexed, reputable publishers/open-access journals, and international peer-reviewed journals in Table 1. Then, semantic network analysis, a powerful and compelling technique in a significant data era, is used to extract patterns or directions with uncovering data-empowered insights. Consequently, 587 unique keywords, 1743 sentences, and 285 paragraphs and documents were identified in 285 abstracts of research articles through NetMiner4’s semantic network analysis program. NetMiner is an eminent application software tool for exploratory analysis and visualization of extensive network data. It also includes several functions and features for data analytics, such as machine learning algorithms, semantic network analysis, etc. It has optimized operations and structure that can efficiently process large amounts of unstructured text data. Accordingly, the study uses this software program, particularly semantic network analysis, which is the most appropriate method to achieve the research questions.

Table 1 The list of journals selected along with selection criteria fulfillment

No	Journal name (Journal quartile)	Data source Platform (publishers’ Web Databases)	Publisher	# of research articles	ISI/ SCOPUS indexed	Reputed publishers/ Open-access journals	International peer-reviewed
1	Active Learning in Higher Education (Q1)	https://journals.sagepub.com/loi/alha	Sage Publications	34	✓	✓	✓
2	Studies in Higher Education (Q1)	https://www.tandfonline.com/toc/cshe20/current	Taylor & Francis	88	✓	✓	✓
3	Teaching in Higher Education (Q1)	https://www.tandfonline.com/toc/cthe20/current	Taylor & Francis	118	✓	✓	✓
4	Internet and Higher Education (Q1)	https://www.sciencedirect.com/journal/the-internet-and-higher-education/issues	Elsevier BV	45	✓	✓	✓
Total			3	285	–	–	–

Source: Scimago (www.scimagojr.com) and each journal website

Data cleaning with refinement

Most commonly, data encloses ambiguous and unnecessary contents. Particularly, unstructured text data include meaningless and illogical messy texts such as symbols, colon, emoticons, consonant and vowel data, and even slang and spelling errors. Thus, it is necessary to filter and purify the data through pre-processing for data cleaning and refinement [5]. In the pre-processed text, the word tab is integrated into the result area. The Word tab provides such information as a list of words, frequency of use, part of speech, which is to be extracted from the unstructured text contents according to the text process settings. After refining, the data are cleaned up, and the quality of the messy data for analysis is improved overall. In this study, messy data were tidy up using Open Refine (used to be called Google Refine), one of the best tools for data cleaning and refinement. Also, the word filtering of data cleaning software in the NetMiner was used for double-checking.

Results and discussion

This study's first objective is to determine the most studied topics in the field of higher education in 2018 and 2019. For this, the top 20 keywords are selected through semantic network analysis among 587 keywords appearing in 285 abstracts of research papers in four top journals related to the issues of teaching and learning in 2018–2019. The result is as shown in Table 2.

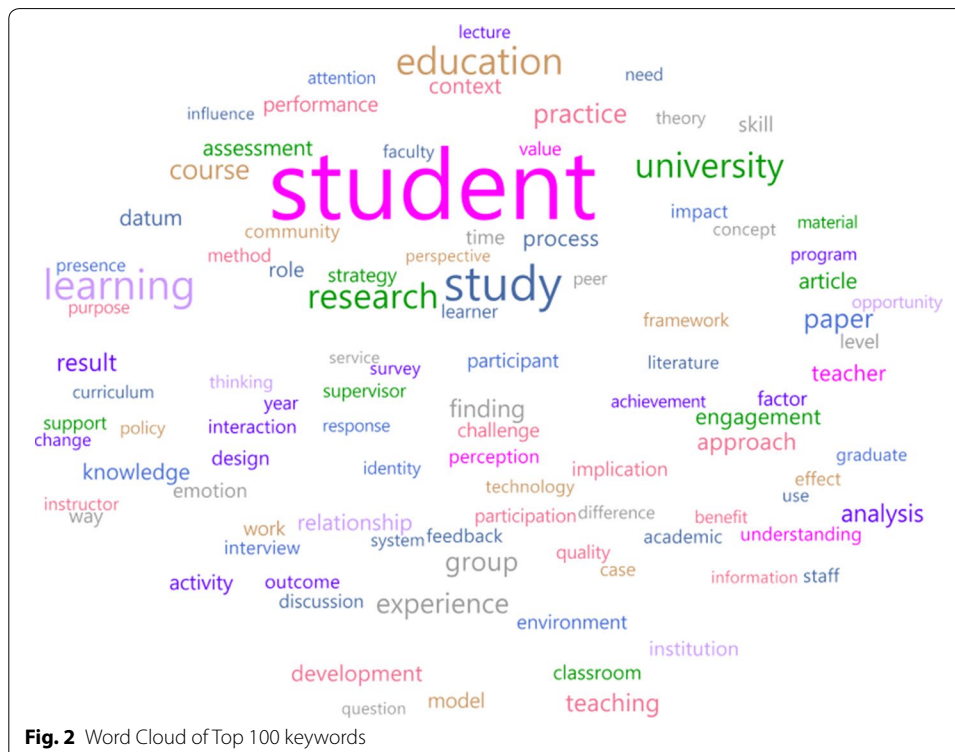
According to the results, the main 'top 20' keywords of teaching and learning in higher education cover the topics of 'students-centered or student-led learning' rather than 'teacher-led' and that, research, experience, group, development, practice, and engagement are identifiable. Especially, the word 'student' composed 7.11% of the total 797 times as a leading keyword showing the highest frequency, and it follows by study, learning, education, research, university, and experience. They are 2.97%, 2.41%, 2.04%, 1.81%, 1.81%, and 1.18%, respectively. In this study, a directional link (Directed Network) based on the 'Binary Network,' which does not weigh the connecting line, is created. Here, the

Table 2 Top 20 key words' term frequency (TF)

Rank	Word	Frequency (%)	In-Degree	Out-Degree	Rank	Word	Frequency (%)	In-Degree	Out-Degree
1	student	797 (7.11%)	144	141	11	practice	124 (1.11%)	20	12
2	study	333 (2.97%)	55	39	12	teaching	103 (0.92%)	18	20
3	learning	270 (2.41%)	33	52	13	result	102 (0.91%)	20	11
4	education	228 (2.04%)	34	28	14	finding	99 (0.88%)	14	13
5	research	203 (1.81%)	40	35	15	analysis	98 (0.87%)	20	11
6	university	203 (1.81%)	33	29	16	approach	94 (0.84%)	6	12
7	experience	132 (1.18%)	12	14	17	process	83 (0.74%)	10	11
8	course	128 (1.14%)	21	20	18	teacher	83 (0.74%)	12	12
9	group	127 (1.13%)	20	19	19	development	81 (0.72%)	8	15
10	paper	126 (1.12%)	21	8	20	engagement	80 (0.71%)	9	14

frequency of a node is an amount defined for each node and means the number of connection lines of each node that exists as a neighbor. In-degree refers to the number of edges coming towards a vertex in a directed graph; out-degree refers to the number of arcs directed away from the vertex. Although Table 2 shows the keywords in the top 20 ranks, it observes that a few keywords in the top and the other keywords show a great difference in the number of nodes in- and out-degree. Accordingly, for the detailed view at a glance, a word cloud is created to check each keyword's quantitative importance. Then, it considers the relationship and features of keywords in more detail through network map and topic modeling. Word Cloud is a visualization tool that illustrates the size of letters according to the importance of keywords. Based on that, we can notice the difference between relatively meaningful words, in brief, to understand how much difference is there. In this study, the word cloud node attributes are used to display information as frequency and number of words by entering the maximum number of words as 100. The result of the word cloud is as follows in the Fig. 2.

Meanwhile, the word-to-word network is visualized with a network map to understand the data analysis results intuitively. It means that the network data is calculated and arranged according to NetMiner's layout algorithm's program. A layout algorithm is a method of calculating where to place nodes to visualize network data. Among the usual methods, Spring Layout is chosen as it can show the connection structure most effectively. Spring layout can be expressed by various algorithms such as Kamada & Kawai, Stress Majorization, and Clustered Eades, Fruchterman & Reingold, GEM, HDE, etc.



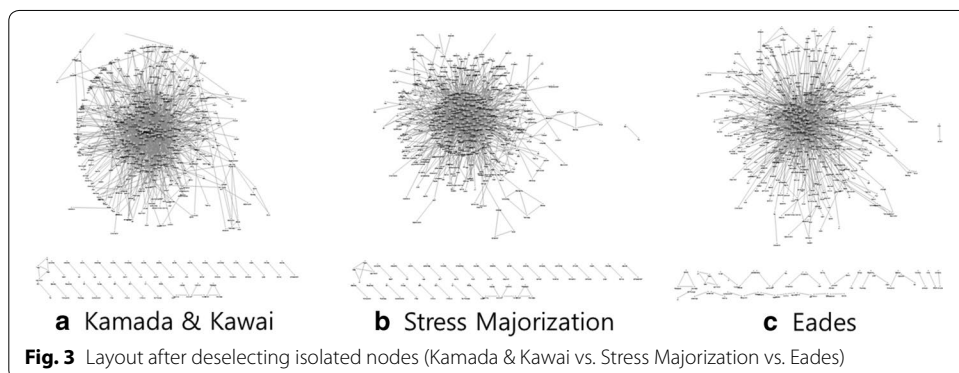
Among them, Kamada & Kawai, Stress Majorization, and Eades are chosen as these algorithms fit the need for data analysis and representation, then compare them as preliminary work to consider the number of various cases of subtopic extraction inherent in a given text. In this study, the isolated nodes are shown at the edge of the network map and visualized comparing the 'Kamada & Kawai,' 'Stress Majorization,' and 'Clustered Eades' network maps deselecting isolated nodes. The layout of the analysis results is as follows in the Fig. 3.

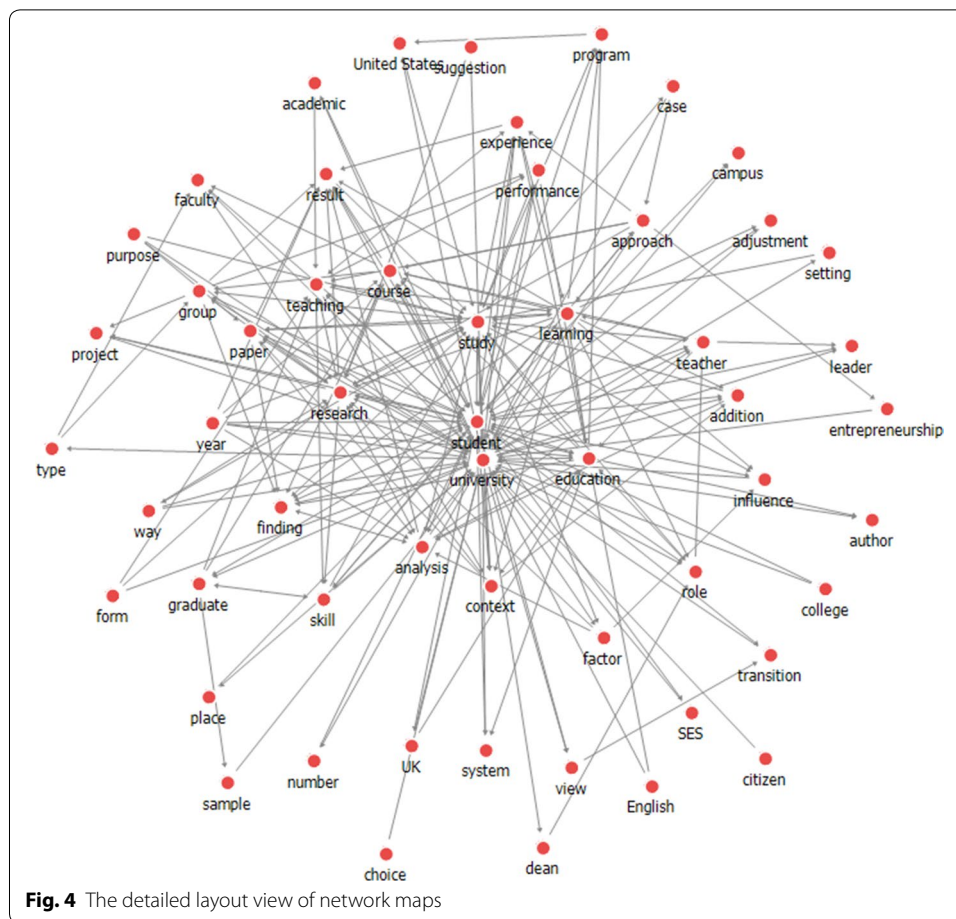
After importing unstructured data, a 2-mode network between words and sentences/paragraphs/documents generated in the program of NetMiner transforms into the keyword-keyword 1-mode system of the research presented in the major journals about teaching and learning. Here's a look at the detailed layout with keywords from the three core types of network maps in Fig. 4.

In the NetMiner program, a 1-mode network generates by using word-to-word distance information. In this study, the nearness of two words is calculated, and based on that, a method of creating links between words located close together is used. The number of words includes setting the link generation range between words called 'Window Size.' In this study, a maximum of 3 words is included in the link generation range by entering Window Size as 3. Links create between words according to the Window Size set as above, and the two linked words are displayed as Source and Target, respectively, in Table 3.

The term TF (Term Frequency) describes above indicates how often a particular word appears in a document. It means the higher the value, the more critical the word. However, a term commonly used (for example, do) may have a high TF value even though it is not an important word. To avoid this, we can measure how many documents appear in a text by Document Frequency (DF). In conclusion, TF-IDF (Inverse Document Frequency) provides information to determine how valuable a word is in a particular document based on word frequency and document frequency. The value of the weight, in general, marks the TF-IDF. This Eq. 1 of the TF-IDF score is calculated as follows.

$$W_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (1)$$





$tf_{i,j}$ = The number of occurrences of i in j ; df_i = The number of documents containing i ; N = total number of documents.

Weight is the link frequency of generated word pairs, meaning that the words of 'student' and 'experience' weight 28, and the word pair appears 28 times in the entire text. # of Sentences/Paragraphs/Documents is the number of sentences/paragraphs/documents in which the word pair appears. In this study, the word pair of 'learning' and 'environment' appear 26 times, with a weight of 26 and # of documents of 19. Gini Coefficient indicates how concentrated the word pair is intensely in a few sentences, paragraphs, and documents. It is also how evenly it appears in multiple sentences, paragraphs, and documents. A value closer to 1 demonstrates that the more focused it is on a few sentences, paragraphs, and documents, the more important the word pair is. In this case, the criterion for sentences, paragraphs, and documents is a co-occurrence unit selected when creating a 1-mode network.

Lastly, the method of LDA (Latent Dirichlet Allocation), which is a machine learning algorithm, is used to extract subtopics embedded in the text [10, 54]. LDA is one of the most popular and influential topic models, a method for analyzing a broad set of unstructured documents. The basic idea is that unstructured documents are represented as a topic distribution where each topic is characterized by a word distribution [12, 54].

Table 3 Top 20 word network

Rank	Source	Target	TF-IDF (Weight)	# of Sentences/ paragraphs/Documents	Gini coefficient
1	student	experience	28	22	0.9
2	learning	environment	26	19	0.9
3	case	study	26	22	0.9
4	student	engagement	25	13	1
5	university	student	25	22	0.9
6	student	perception	24	16	1
7	student	learning	22	18	0.9
8	study	student	22	22	0.9
9	learning	student	20	18	0.9
10	education	institution	17	17	0.9
11	teaching	learning	16	12	1
12	student	course	16	13	1
13	student	university	16	14	1
14	education	student	15	15	0.9
15	experience	student	13	13	1
16	student	staff	12	9	1
17	learning	experience	11	11	1
18	learning	process	11	9	1
19	student	teacher	11	6	1
20	focus	group	11	11	1

We can denote $p(z|d_i)$, $p(w|z_{i,j})$ as the topic distribution for each document i , and the word distribution for a topic allocated to j th word of document i , respectively. In the learning phase, LDA fits $p(z|d_i)$, $p(w|z_{i,j})$ to a pair of documents (i.e., a document-by-word sparse matrix). Given these distributions, the LDA can generate a new document with the following generative process [12]:

for j th word in the i th document:

Choose a topic $z_{i,j} \sim \text{Multinomial}(p(z|d_i))$

Choose a word $w_{i,j} \sim \text{Multinomial}(p(w|z_{i,j}))$.

For each topic, the top nodes from the main nodes show in Table 4. Moreover, this table shows the number of nodes and probability included for each topic when the classification labels of nodes in a Subnodeset are assigned to the topic that maximizes the topic proportion from ‘SubNode’.

MainNode (it is a Keyword) and SubNode (it is a Document) show probability information about which topic to associate. Using this method, we can figure out that the more similar the values of the response variables are, the more likely they are to contain compatible subjects. In contrast, the more significant the differences between them, the more likely they are to include different topics. The results of the top 3-topic modeling, according to the # of documents, show as follows. The main subject of the first topic can be interpreted as that practice and learning assessment can be achieved through various activities such as community activities. The second topic suggests that students’ college education can also reach through experience-based classes as staff. Third, the supervisor’s research may indicate that subjects in academic contexts can be addressed. In addition to that, several issues are revealed as uprising topics through the topic modeling of

Table 4 Topic modeling using LDA (Latent Dirichlet allocation model)

	1st Keyword (Prob.)	2nd Keyword (Prob.)	3rd Keyword (Prob.)	4th Keyword (Prob.)	5th Keyword (Prob.)	# of documents
Topic-1	learning (0.113)	practice (0.037)	assessment (0.034)	activity (0.028)	community (0.023)	226
Topic-2	student (0.149)	university (0.085)	experience (0.049)	classroom (0.016)	staff (0.01)	199
Topic-3	research (0.099)	paper (0.045)	supervisor (0.023)	context (0.023)	academic (0.02)	198
Topic-4	education (0.114)	approach (0.036)	institution (0.019)	system (0.014)	study (0.012)	188
Topic-5	group (0.071)	result (0.038)	participant (0.024)	student (0.023)	effect (0.016)	182
Topic-6	skill (0.039)	knowledge (0.037)	model (0.029)	performance (0.027)	role (0.026)	174
Topic-7	student (0.09)	study (0.082)	emotion (0.032)	challenge (0.025)	perception (0.021)	164
Topic-8	student (0.085)	analysis (0.035)	relationship (0.032)	strategy (0.025)	feedback (0.024)	154
Topic-9	student (0.073)	study (0.066)	course (0.048)	outcome (0.026)	case (0.019)	132
Topic-10	teaching (0.048)	finding (0.04)	teacher (0.034)	student (0.028)	practice (0.027)	126

LDA. For example, a system-based education, the results or effects of group activity, how skills and knowledge can play a critical role in performance models. Moreover, how students' perceptions or feelings can affect learning outcomes, relationships with students, and feedback could be analyzed to develop learning strategies, learning outcomes using cases, and finding teaching strategies through students' and teachers' practices.

Implication and conclusion

This study aspires to grasp the latest research agendas and academic trends in teaching and learning with the keywords of major international journals in higher education through semantic network analysis. Consequently, it turns out that this study obviously provides educators, researchers, and academic leaders with data-empowered insights and intuition. The identifiable future agendas, such as a trend-based approach in teaching, learning, and research in higher education, will guide future education with right direction. In this regard, this study outlines the significant implications as follows.

First of all, the world is changing faster than ever; accordingly, it is not easy to follow up on the current educational trends in higher education by mastering whole agendas. Most academic members do not adapt as quickly to keep them up-to-date in the double loop with strategic agility. Given that, this study allows us to identify the current flow of education and educational issues at a glance by leveraging advanced technology and big data. This study focuses on the most recent 2 years of big research data to best support this current situation rather than looking at old panel data such as ten or twenty-year past data. Recent big data generated by reputable publishers' web database platforms and analytics has become an essential component for the higher education sector. In this way, this research reflects best the current educational situations

in teaching and learning of higher education and contribute to preparing for the future with data-empowered intuitions and insights.

Secondly, the semantic network analysis method, a data-driven approach used in this paper, may shed light on the development of new techniques using the machine learning algorithm to get the whole picture of a new paradigm or shift in higher education. It proves that this is a very effective and useful tool to understand the key attributes of current flow and network map in teaching and learning thoroughly by extracting main issues and associated sub-topics. In this regard, this study enhances a more objective view of the actual educational reality and phenomena through big data and machine learning models in the educational sector entirely.

Lastly, traditionally, 'teacher-centered' or 'teacher-led' education in teaching and learning was the central theme in higher education. A teacher or instructor mainly focused on delivering entire contents from his or her side to students unilaterally. However, now, the paradigm is wholly changed into 'student-centered' or 'student-led' education. Then, various issues, such as learning, education, research, and experience that follow incidentally, are broken down and spread to diverse subjects. This study firmly reflects in this modern paradigm and educational trend quantitatively and qualitatively. Furthermore, this research result will play a critical role in reshaping educational environments and critical perspectives on teaching and learning in higher education placed on big data.

Although this research suggests several substantial and critical implications, there are still some limitations to be improved below. It will address in future research.

In this study, 285 scholarly articles from the top-ranked international journals related to teaching and learning in 2018–2019 were gathered and analyzed to recognize the most recent research agendas and trends, as it intended, intends to see only up-to-date data given that the trend is changing quickly. However, it is a fact that studies using massive data improve the results in terms of the level of validity and reliability of analysis in general, and it still works in the academic world. Accordingly, future research is needed to collect more data by increasing the number of articles on teaching and learning.

Next, the semantic network analysis or semantic mining technique performs statistical, logical, and rule-based semantic networks analysis quantitatively and qualitatively. However, the research method still leaves room for consideration. Traditional network analysis techniques analyzed physical world relationships simply based on distance (such as centrality), strength, direction, etc. However, recently network analysis has been developed continuously. The giant network is continually forming and flowing subsequent knowledge as large as 1: N link for real-time communication through social media such as Snapchat, Instagram, and YouTube, etc. Thus, it should notice that recent network analysis is very complicated, like the social network web, with a vast amount of data flowing through the network structure. Also, it is necessary to derive and analyze sub-networks aiming to apply them to diverse sectors by reflecting those attributes of a network into the research world.

Lastly, approaching individually or together with a holistic perspective in different education sectors such as law, engineering, business, computer science, or any other studies in higher education is worth doing for future research. The discovery and comparison of the most recent issues or keywords on teaching and learning in each field of higher education will help drive innovation and change in teaching and learning entirely.

Furthermore, it encourages academic leaders, more senior or higher management teams, to have effective and proactive leadership by realizing the trend-based education and following up the future directions.

Acknowledgements

Not applicable.

Authors' contributions

The author confirms the sole responsibility for this manuscript fully as a sole author for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation. The author read and approved the final manuscript.

Funding

Not applicable. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Availability of data and materials

In this study, the data supporting the findings are openly available in publishers' web platforms of four journals, respectively. These data are derived from the articles of each journal in the public domain.

Competing interests

The author reports no potential conflict of interest.

Received: 30 July 2020 Accepted: 8 October 2020

Published online: 21 October 2020

References

- Algarni A. Data mining in education. *Int J Adv Comput Sci Appl*. 2016;7(6):456–61. <https://doi.org/10.14569/ijacs.a.2016.070659>.
- Angelo TA. Transforming assessment: high standards for higher learning. *AAHE Bulletin*. 1996;48:3–4.
- Atteveldt WV. Semantic network analysis: techniques for extracting, representing, and querying media content. Charleston: BookSurge Publishers; 2008.
- Ausubel DP. Educational psychology: a cognitive view. London: Holt Rhinehart & Winston; 1968.
- Balan S, Rege J. Mining for social media: usage patterns of small businesses. *Bus Syst Res*. 2017;8(1):43–50.
- Banta T, Land J, Black K, Olander F. Assessment in practice: Putting principles to work on college campuses. San Francisco: Jossey - Bass; 1996.
- Barrie SC, Prosser M. Aligning research on student learning with institutional policies and practices on evaluation and quality assurance, Paper presented at the 11th ISL Conference, Brussels, 4–6 September; 2002.
- Bauer M, Henkel M. Responses of academe to quality reforms in higher education: a comparative study of England and Sweden. *Tert Educ Manag*. 1997;3(3):211–28.
- Biggs J. The reflective institution: assuring and enhancing the quality of teaching and learning. *High Educ*. 2001;41(3):221–38.
- Blei D, Ng A, Jordan M. Latent dirichlet allocation. *J Mach Learn Res*. 2003;3:993–1022.
- Bose R. Advanced analytics: opportunities and challenges. *Ind Manag Data Syst*. 2009;109(2):155–72. <https://doi.org/10.1108/02635570910930073>.
- Cyram NetMiner (2019). NetMiner Semantic Network Analysis Manual. Cyram. Retrieved from <https://www.NetMiner.Com>.
- Deng R, Benckendorff P, Gannaway D. Progress and new directions for teaching and learning in MOOCs. *Comput Educ*. 2019;129:48–60.
- Drieger P. Semantic network analysis as a method for visual text analytics. 9th Conference on Applications of Social Network Analysis (ASNA). *Procedia Soc Behav Sci*. 2013;79:4–17. <https://doi.org/10.1016/j.sbspro.2013.05.053>.
- Doerfel ML. What constitutes semantic network analysis? A comparison of research and methodologies. *Connections*. 1998;21(2):16–26.
- Elton LRB, Laurillard DM. Trends in research on student learning. *Stud High Educ*. 1979;4(1):87–102. <https://doi.org/10.1080/03075077912331377131>.
- Entwistle NJ. Complementary paradigms for research and development work in higher education, Conference of the European Association for Research and Development in Higher Education, Rotterdam; 1973.
- Executive Core. Future Trends in Business Education. Summer 2015. Executive Core, LLC. 2015; https://www.uniconexed.org/wp-content/uploads/2017/01/Future_Trends-Revised_Final_Report_Executive_Core-August-Jan.12.2016-v2.pdf.
- Foster J. Key Trends Influencing Graduate Business Education in 2019. <https://www.mba.com/article-and-announcements/articles/why-business-school/key-trends-influencing-graduate-business-education-in-2019>; 2019.
- Guri-Rozenblit S. Distance/Open Learning-trends and developments as reflected in recent literature. *Stud High Educ*. 1991;16(1):83–90.
- Heilstra TM, Siguroardottir MS. The flipped classroom: does viewing the recordings matter? *Act Learn High Educ*. 2017;19(3):211–23. <https://doi.org/10.1177/1469787417723217>.
- Henderson M, Selwyn N, Aston R. What works and why? Student perceptions of 'useful' digital technology in university teaching and learning. *Stud High Educ*. 2017;42(8):1567–79. <https://doi.org/10.1080/03075079.2015.1007946>.

23. Iniguez, S. (2015). Major trends in business education. *MBA Journal*. 2015 12:16. <https://www.mbjournal.de/major-trends-in-business-education/>.
24. Jarvenoja H, Naykki P, Tormanen T. Emotional regulation in collaborative learning: when do higher education students activate group level regulation in the face of challenges? *Stud High Educ*. 2019;44(10):1747–57. <https://doi.org/10.1080/03075079.2019.1665318>.
25. Kelly E, O'Connell P, Smyth E. The economic returns to field of study and competencies among higher education graduates in Ireland. *Econ Educ Rev*. 2010;29(4):650–7. <https://doi.org/10.1016/j.econedurev.2009.11.001>.
26. Kharlamov A, Gradoselskaya G, Dokuka S. Dynamic semantic network analysis of unstructured text corpora. *International Conference on Analysis of Images, Social Networks and Texts (AIST 2017)*, p. 392–403; 2017.
27. Kim HM. Analysis of research trends of South Korean music education through semantic network analysis. *Korean J Res Music Educ*. 2015;44(4):49–68.
28. Kim ND, SNU Consumer Trend Analysis Center. *Trends Korea 2020. Futuristic window (Milaeui Chang)*: Seoul, Korea; 2019.
29. Kim M, Choi M, Youm Y. Semantic network analysis of online news and social media text related to comprehensive nursing care service. *J Korean Acad Nurs*. 2017;47(6):806–16. <https://doi.org/10.4040/jkan.2017.47.6.806>.
30. Lalley C, Houston J, Gasteen A. Gender disparities in horizontal mismatch penalties: an examination of professional degrees in the UK (2007–2015). *Stud High Educ*. 2018;44(12):2265–80. <https://doi.org/10.1080/03075079.2018.1485145>.
31. Lee SJ. Research trends of Korean journalism and communication studies using a semantic network analysis. *J Korea Contents Assoc*. 2016;16(7):179–89.
32. Lee S, Choi JH, Kim HW. Semantic network analysis on the MIS research keywords: APJIS and MIS Quarterly 2005–2009. *Asian Pac J Inf Syst*. 2010;20(4):25–51.
33. Levine D. Trends in Technology: How we work, live and consume. *USA Trends Day Keynote*. 2019; <https://www.danillevine.com/>.
34. Lomer S, Anthony-Okeke L. Ethically engaging international students: student generated material in an active blended learning model. *Teach High Educ*. 2019;24(5):613–32. <https://doi.org/10.1080/13562517.2019.1617264>.
35. Marginson S, Considine M. *The enterprise university: Power, governance and reinvention in Australia*. Cambridge: Cambridge University Press; 2000.
36. Martin F, Wang C, Sadaf A. Student perception of helpfulness of facilitation strategies that enhance instructor presence, connectedness, engagement and learning in online courses. *The Internet and Higher Education*. 2018;37:52–65. <https://doi.org/10.1016/j.iheduc.2018.01.003>.
37. Maxwell S. *Top 3 MBA Trends in 2019*. *Hiperpool*. 2019, <https://hiperpool.com/blog/top-3-mba-trends>.
38. Miner G, Elder J, Fast A, Hill T, Nisbet R. *Practical text mining and statistical analysis for non-structured text data applications*. 1st ed. Cambridge: Academic Press; 2012.
39. Mishra M, Mishra P, Somani AK. Understanding the data science behind business analytics (Chapter 5, p. 93–116). In: *Big data analytics: tools and technology for effective planning*. CRC Press. 2017. <https://doi.org/10.1201/b21822>.
40. Monge PR, Eisenberg EM. Emergent communication networks. In: Jablin FM, Putnam LL, Roberts KH, Porter LW, editors. *Handbook of organizational communication: An interdisciplinary perspective*. Thousand Oaks: Sage Publications Inc; 1987. p. 304–42.
41. Nikitina T, Lapina I. Overview of Trends and Developments in Business Education, The 21st World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2017, Vol.2, USA, Orlando. 2017.
42. Nomuoja JO (2010) Current trends in business education in higher institutions. www.globalacademicgroup.com.
43. Nulty, P. (2017). Semantic Network Analysis of Contested Political Concepts. *International Conference on Computational Semantics (IWCS 2017)*, Retrieved from <http://www.aclweb.org/anthology/W/W17/#6800>.
44. OECD. *Trends Shaping Education 2019 Spotlight 18 Play!*. Paris: OEDC Publishing; 2019.
45. Park SY. An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-Learning. *Educ Technol Soc*. 2009;12(3):150–62.
46. Park YE, Alenezi M. Predicting the popularity of Saudi multinational enterprises using a data mining technique. *J Manag Inf Decis Sci*. 2018;21(1):1–14.
47. Park YE. Data Empowered Insights for Sustainability of Korean MNEs. *J Asian Financ Econ Bus*. 2019;6(3):173–83. <https://doi.org/10.13106/jafeb.2019.vol6.no3.173>.
48. Pitan OS, Muller C. University reputation and undergraduates' self-perceived employability: mediating influence of experiential learning activities. *High Educ Res Dev*. 2019;38(6):1269–84. <https://doi.org/10.1080/07294360.2019.1634678>.
49. Price L, Kirkwood A. Using technology for teaching and learning in higher education: a critical review of the role of evidence in informing practice. *Higher Education Research and Development*. 2014;33(3):549–64. <https://doi.org/10.1080/07294360.2013.841643>.
50. Rice RE, Danowski JA. Comparing comments and semantic networks about voicemail. *Proc Am Soc Inf Sci*. 1991;28:134–8.
51. Shen CW, Ho JT. Technology-enhanced learning in higher education: a bibliometric analysis with latent semantic approach. *Comput Hum Behav*. 2020. <https://doi.org/10.1016/j.chb.2019.106177>.
52. Shneiderman B, Aris A. Network visualization by semantic substrates. *IEEE Trans Visual Comput Graphics*. 2006;12(5):733–40.
53. Srivastava J, Srivastava AK. Data mining in education sector: a review. *Special Conference Issue: National Conference on Cloud Computing & Big Data*. 184–190. 2014. <https://www.ijana.in/Special%20Issue/C33.pdf>.
54. Steyvers M, Griffiths T. Probabilistic topic models. In: Landauer TK, McNamara DS, Dennis S, Kintsch W, editors. *Handbook of latent semantic analysis*. New Jersey: Lawrence Erlbaum Associates Publishers; 2007. p. 427–48.
55. Steyvers M, Tenenbaum JB. The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. *Cognit Sci*. 2005;29(1):41–78.
56. Stohl C. European managers' interpretations of participation: a semantic network analysis. *Hum Commun Res*. 1993;20(1):97–117. <https://doi.org/10.1111/j.1468-2958.1993.tb00317.x>

57. Strauss H, Maisonneuve CDL. The Wage Premium on Tertiary Education: new Estimates for 21 OECD Countries. *OECD J Econ Stud*. 2009;2009:183–210.
58. Tierney A. The scholarship of teaching and learning and pedagogic research within the disciplines: should it be included in the research excellence framework? *Stud High Educ*. 2020;45(1):176–86. <https://doi.org/10.1080/03075079.2019.1574732>.
59. Venkatesh R, Balasubramanian C, Kaliappan M. Development of big data predictive analytics model for disease prediction using machine learning technique. *J Med Syst*. 2019;43(8):272. <https://doi.org/10.1007/s10916-019-1398-y>.
60. Wood A, Galloway RK, Sinclair C, Hardy J. Teacher-student discourse in active learning lectures: case studies from undergraduate physics. *Teach High Educ*. 2018;23(7):818–34. <https://doi.org/10.1080/13562517.2017.1421630>.
61. Wright J. Methodology review, part 2: text mining and semantic network analysis. 2018. <https://jaredmwr.wordpress.com/2018/03/13/methodology-review-part-2-text-mining-and-semantic-network-analysis/>.
62. Xing W, Tang H, Pei B. Beyond positive and negative emotions: looking into the role of achievement emotions in discussion forums of MOOCs. *Internet High Educ*. 2019. <https://doi.org/10.1016/j.iheduc.2019.100690>.
63. Yun E, Park Y. Extraction of scientific semantic networks from science textbooks and comparison with science teachers' spoken language by text network analysis. *Int J Sci Educ*. 2018;40(17):2118–36. <https://doi.org/10.1080/09500693.2018.1521536>.
64. Zainuddin Z, Halili SH. Flipped classroom research and trends from different fields of study. *Int Rev Res Open Distance Learn*. 2016;17(3):313–40.
65. Zaki MJ, Meira W. *Data mining and analysis: fundamental concepts and algorithms*. London: Cambridge University Press; 2014.

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
