electronics

Article

# A Sentence-Embedding-Based Dashboard to Support Teacher Analysis of Learner Concept Maps

**Filippo Sciarrone** [1,*,†] and **Marco Temperini** [2,*,†]

[1] Faculty of Technological and Innovation Sciences, Universitas Mercatorum, 00186 Rome, Italy
[2] Department of Computer Control and Management Engineering, Sapienza University of Rome, 00189 Rome, Italy
* Correspondence: filippo.sciarrone@unimercatorum.it (F.S.); marte@diag.uniroma1.it (M.T.)
† These authors contributed equally to this work.

**Abstract:** Concept mapping is a valuable method to represent a domain of knowledge, also with the aim of supporting educational needs. Students are called upon to construct their own knowledge through a meaningful learning process, linking new concepts to concepts they have already learned, i.e., connecting new knowledge to knowledge they already possess. Moreover, the particular graphic form of a concept map makes it easy for the teacher to construct and interpret both. Consequently, for an educator, the ability to assess concept maps offered by students, facilitated by an automated system, can prove invaluable. This becomes even more apparent in educational settings where there is a large number of students, such as in Massive Open Online Courses. Here, we propose two new measures devised to evaluate the similarity between concept maps based on two deep-learning embedding models: InferSent and Universal Sentence Encoder. An experimental evaluation with a sample of teachers confirms the validity of one such deep-learning model as the baseline of the new similarity measure. Subsequently, we present a proof-of-concept dashboard where the measures are used to encode a concept map in a 2D space point, with the aim of helping teachers monitor students' concept-mapping activity.

**Keywords:** concept map; learning analytics; deep learning



**Citation:** Sciarrone, F.; Temperini, M. A Sentence-Embedding-Based Dashboard to Support Teacher Analysis of Learner Concept Maps. *Electronics* **2024**, *13*, 1756. https://doi.org/10.3390/electronics13091756

Academic Editor: Ioannis Hatzilygeroudis

Received: 27 February 2024
Revised: 14 April 2024
Accepted: 25 April 2024
Published: 2 May 2024



## 1. Motivation and Goals

The continuing surging ahead of digital technologies and their applications is at the basis of advances in all sectors of human activity, and in particular, according to the topics of interest in this paper, in Education [1–5].

Learning, teaching, and training activities, in general, are greatly helped if one can make use of tools supporting the shareable representation of one's knowledge as it is possessed and organized in one's mind. Concept Maps (CMs) are a well-known method of reaching the above aim. CMs are increasingly adopted to support teaching and learning [6], with aims such as to represent the mental organization of knowledge in a learner, to unveil misconceptions, or to allow teachers to design a course, build it, and provide learners with a visual demonstration of the course's knowledge domain [7,8], or to unveil misconceptions in learners, and produce assessments [9].

More in general, the "concept" itself of concept mapping is a tool that allows for the representation and analysis of how humans construct and organize knowledge. Differently than other formal methods used to visualize knowledge elements and their relations, in CMs, there is not a predetermined semantics for the words (the vocabulary) used in a CM, or at least such a semantic is not associated uniquely with the CM. This characteristic has a two-fold effect on a CM tool. On the one hand, it is, in fact, an advantage over more formal representations (such as the one in semantic nets or ontologies), allowing for use in informal environments (as a classroom can easily be). On the other hand, vocabulary freedom implies a great deal of effort when the CMs must undergo an automated process of analysis and management. Regarding the second point, we must take into account that such automated management is going to be increasingly unavoidable as the wealth of electronically defined and stored CMs is exponentially increasing.

Actually, having the possibility to use a software system able to assess the similarity of a CM to another (from a topological and semantic viewpoint) can prove highly valuable for more reasons than one: (1) it could enable educators to gain insights into students' learning trajectories, including the related temporal progression; (2) educators can establish repositories of CMs, beneficial for initiating pedagogical initiatives such as producing a fresh course or a new set of learning activities to explore a novel domain of knowledge; (3) educators could categorize their students based on the *learning distance* existing among them, based on their different proficiency within a specific domain of knowledge.

Research works on the comparison of CMs are numerous. Many of the proposed methods rely on the maps' topological structure, represented as a Directed Acyclic Graph (DAG), and/or on the educational arrangement of concepts and relationships in the maps [8]. Nevertheless, it is worth noting that the perceived similarity between two CMs can vary depending on the observer's perspective and on the possibly different aspects that are deemed significant for the CM similarity. Examples of different aspects might be the simple commonality of nodes (concepts), which would disregard the context of their appearance in the maps, or we might privilege the fact that the same concepts in the two CMs have (more or less) the same prerequisite knowledge, as a consequence of their position in the maps; alternatively, the cognitive load imposed by the concepts could be primarily considered.

A global, general method to define the similarity between two CMs should indeed result from the weighted combination of different measures of similarity, also according to the teacher's preferences.

In this work, we propose the application of certain deep-learning (DL) models to analyze the CMs produced by a class of learners. The analysis consists of determining the relations of similarity among the maps from a semantic, cognitive point of view. Such analysis is conducted using a measure of similarity we define and present here.

In short, our approach to similarity computation works on a specific representation of each individual CM: first, the CM is transformed in a set of sentences (propositions); then, each proposition is represented in an *embedding*, i.e., a fixed-length vector of real values, and then the CM is considered to be represented by the set of such embeddings.

The measure of similarity between two CMs can then be computed as the distance between the two maps' sets of embeddings through methods such as the *Euclidean distance*, or the *Cosine-similarity distance*. In this article, we use both such methods for corroboration.

In particular, as embedding methods, in the experimental work we present here, we used two DL models: the *InferSent* model [10], and the *Universal Sentence Encoder* [11].

An initial assessment of the measure of similarity, outlined in [12], proved satisfactory. Consequently, in this paper, we suggest utilizing this similarity to aid the entire educational process. Specifically, we outline the definition of a Learning (and Teaching) Analytics process, utilizing our measure of similarity, also showing its implementation in a six-step didactic process, where the teacher is supported by a dashboard [13,14] to monitor and improve the concept-mapping activity. In our intentions, the process and the measure would provide the teacher with insights on the level of appropriation of knowledge and could help both unveil misconceptions and appreciate non-straightforward connections in the learner's map.

In terms of research results, in this work, we wish to present, basically, two results: one is associated with the Research Question (henceforth RQ) defined in the next paragraph, and its answer comes from the analysis of experimental activity involving several teachers:

**RQ** Can a similarity measure be defined through the deep-learning techniques we selected so as to be able to capture the commonalities between two CMs in a suitable and useful manner, from a pedagogical viewpoint, for the teacher?

The other result, in our view, is in the design of a teacher dashboard. We present a proof-of-concept dashboard with no intention of providing it as a complete system to be used right away in the classroom. Rather, we show it to let the reader see how it can be useful to a teacher for the analysis of the CMs produced by her/his (possibly large) class. After having described a process of Learning and Teaching Analytics designed to suit the educational use of CMs, we describe the dashboard, which is intended to support such a process, by presenting data elaborated through our sentence-embedding-based CMs similarity measure. In particular, we show how the dashboard can depict relevant data about the maps, i.e., all the class' many maps, for the teacher. We present the dashboard here, in its prototypical implementation, as a concept system to show how the graphical rendering can help the teacher's analysis work while demonstrating a possible application of our similarity measure.

In our ambition, the dashboard can allow the teacher to perform a preliminary analysis of the set of CMs produced by the class as answers to an assignment. The aim of the preliminary analysis would be the evaluation of how the class currently understands the knowledge domain involved in the task, allowing the teacher to undertake important educational actions, such as the assessment of the learning activity so far undergone, the comprehension of difficulties existing among the students, the improvement of the learning material, the uncovering of misconceptions, more or less frequent as they might be. One interesting aspect here is also the possibility of performing such analysis in a formative assessment context, where the fear of a bad grade would less influence the students' behavior. All the above are, in fact, significant educational challenges: to meet them in a technologically supported context, where the pressure of the grade is lighter, and the dashboard would lighten the analysis work, can prove to be a significant help for the teacher, and in turn for the students.

In the following, Section 2 briefly introduces some important works on CM similarity measures, together with a short background on CMs. Then, Section 3 will describe some embedding techniques applied to CMs and define our method. In Section 4, the experimental data are presented, and the RQ is discussed, while in Section 5, the implementation of the dashboard system is given. Section 6 discusses the RQ and the expected usefulness of the teacher's dashboard. Finally, in Section 7, some conclusions are drawn.

## 2. Background and Related Work

To show the context of our research, in this section, we present in short some interesting contributions with which we came into contact. Basically, there are two topics we would like to discuss: on the one hand, the nature, and use of CMs; on the other hand, the efforts made in relation to the definition of measures, or qualitative and quantitative methods, devised to support the analysis of CMs, and in particular the analysis of similarity between maps.

### *2.1. Concept Maps*

J. Novak introduced CMs as a strategy to support students in learning scientific concepts (and represent their meaning in an organized, shareable fashion, useful for discussion and deepening) [15]. The theory is based on *Cognitive Constructivism*. Among the pedagogical concepts behind such theory are *Meaningful Learning*, and the consideration of prior knowledge's role in learning [16], scaffolding management, and the idea of learning by experience[17], and cognitive schemes theory, and the learning foundations in social interaction, with the *zone of proximal development* [18].

Basically, a CM visually represents a domain of knowledge, where *concepts* and *relationships* between concepts are depicted concisely as *nodes*, containing the words representing concepts, connected by *edges*. An edge is a line labeled by a verbal expression, connecting a node to a successor node and representing the relationship semantically linking a node to the other. The sequential connection of more nodes through successive edges depicts the paths of connected nodes. Cross-concept relationships, i.e., edges going from one concept in a path to another in another path, can also be drawn. Thus, a CM offers a graphical and symbolic representation of a knowledge domain, with concepts appearing earlier, when more abstract, or later, when more specific, along the paths. Figure 1 shows an illustration of a CM on the *Molecules* domain: the primary concept resides in the top node while connections between pairs of concepts are depicted by lines, each labeled according to the relationship type. Furthermore, each line denotes a triplet: *head concept* (h), *relationship* (r), *target concept* (t). Later, we will show how every triplet forms a sentence that can be subsequently linked to others via new relationships. Consequently, a CM can be interpreted as a collection of sentences (which we will call $C^S$) and progresses through hierarchical levels of depth.

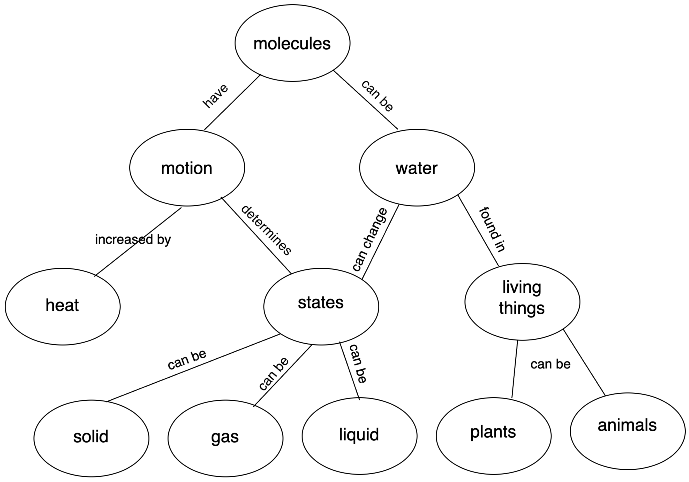


**Figure 1.** A representation of *Molecule* Knowledge domain through a CM.

The idea of CMs stems from investigations into memory. Among various topics, such studies tackle the issue of *Rote Learning* (i.e., in short, memorization) versus *Meaningful Learning* (in very short, understanding and being able to build on it): an individual's knowledge is not merely an assortment of ideas; it is a network of interconnected concepts [16,19]. During Meaningful Learning, it is indeed this network that undergoes alteration, with the introduction of new ideas, the establishment of new connections among them, and potentially the formation of new connections between new and existing ideas. The visual representation of a knowledge domain, with the visualization of both connections (edges) and concepts (nodes), can benefit both educators and learners. When a knowledge representation is provided for the domain, educators can access/visualize a learner's knowledge state, while learners can effectively display their knowledge and gain assistance in comprehending and learning about the domain [15,19]. Thus, a CM serves as a visual and individualized depiction of an individual's knowledge regarding a specific topic [15]. For a (subset of a) knowledge domain, each learner will construct a distinct CM, furnishing valuable insights into their learning process within that domain. Furthermore, the capability to gauge the similarity between two distinct CMs, whether constructed by different students simultaneously or by the same student at various steps in their learning process, would aid educators in monitoring the students' learning process (and assist in uncovering instances of Meaningful Learning [16].

### *2.2. Concept Maps Tools*

Numerous cloud-based web platforms enable the creation and dissemination of CMs using visual techniques like *drag and drop*.

The *Graphed* framework [20] offers users a visual web-based environment for constructing CMs. This platform provides users with a selection of standard relationships between concepts to construct their CMs. Additionally, *Graphed*-managed CMs can be made accessible and shared for the community of users (https://graphed.igiresearch.com/ accessed on 20 February 2024).

The web-based system *Cmaptools* [21] assists users in constructing, navigating, sharing, and discussing domains depicted as CMs. CMs can be shared publicly or in groups of users. Additionally, users can establish links between their CMs and shared CMs and ask for an automated generation of a web page for their CMs. Lastly, users can search the web for information relevant to a CM (https://cmap.ihmc.us/cmaptools/ accessed on 23 February 2024).

We think we should mention, here, *ChatGPT* (https://www.openai.com accessed on 13 March 2024), as a state-of-the-art intelligent chatbot based on a Large Language Model architecture. It can be integrated into learning systems via an API and can provide support for CM construction. In particular, the system can provide CMs both in text and graphic format on demand, based on the user's textual specifications. On the other hand, even the most advanced version, ChatGPT 4.0, is currently unable to provide, off the shelf, direct and graphical support to a teacher based on a reasonable measure of similarity between maps. Basically, this is due to the lack of a well-founded measure of similarity that the system could find available. It also may depend on the lack of image management, which could, anyway, be a problem that is open to solution by a textual specification of the maps.

In conclusion, the aforementioned systems, while mostly providing a good graphical and interactive environment for the construction of CMs, do not offer any measure of distance between them.

In the following subsection, we discuss further research involving in some way the use of measures of similarity, or distance, between maps. This is to better explain the novelty the authors think this article's contribution can provide.

### *2.3. Measures and Analysis Methods for Concept Maps*

A literature review about the use of open-ended CMs, also for educational aims, is presented in [6]. In particular, the comparison of literature contributions is focused on how maps are analyzed, on which maps' characteristics the analysis is based, and to what conclusions, about the learner, an analysis leads. Open-ended CMs are those created by learners when there are no constraints on the vocabulary of concepts' and relationships' names (as opposed to the assignment in which names are given, for instance, by the teacher). The above freedom allows the consideration of the map's true representations of the state of knowledge of the learner who created the map [22]; this is the reason this type of concept mapping is preferably used for monitoring and assessment [23].

An investigation of the possible methods of CM analysis, including the consideration of both the topological structure and morphology of a map, is presented in [7]. In this paper, the topology of the map is given, quite traditionally, by the ways the nodes (concepts) are connected, without considering the semantics of such nodes and connections: this allows for general considerations about the number of concepts and the distribution, and number, of the connections among the nodes, from a structural viewpoint. The morphology instead expresses the graphic form of the map, which can be related to the level of appropriation of knowledge and the depth (level of possession) of such knowledge in the learner's mind. The morphological viewpoint is related to qualities of the map such as *broad* (rather than deep), *deep* (rather than broad), *balanced* (in the paper, one concept is considered "foundational", with all the other concepts deriving from it, in a tree-like form), *disconnected* (with separated "islands" in the map), *interconnected* (with connections among parts that otherwise would be disconnected). The paper proposes some ideas about how to make the two-level, quantitative and qualitative, analysis fruitful to better understand the learner's state of knowledge. An interesting citation proposed by the authors should be cited here: They report an observation from [24] about CMs that result in networks (as opposed to the easier chain of concepts); such maps represent *a rich body of knowledge in which complex understanding is demonstrated*. The paper follows up on this, adding a later observation from [25] amounting to affirm that the *expertize* [shown by a learner] is the *ability to dynamically transform between 'chains of practice and the underlying networks of understanding'*. This, in turn, relates to a note in [26], expressing the idea that *learning involves a transition between 'written or spoken messages [that] are necessarily linear sequences of concepts and propositions' and 'knowledge [which] is stored in our minds in a kind of hierarchical or holographic structure'*.

In [27], the authors investigate the limitations inherent in similarity measures of relatedness proximity between concept pairs in a CM. To this aim, they present a research model for augmenting CMs based on a novel method of co-word analysis that utilizes webometrics web counts for improving similarity measures. The proposed experiment demonstrates the effectiveness of this method in producing reliable and weighted relationships between different concepts in a CM.

A bouquet of similarity measures between CMs is proposed in [8]. The (three) similarity measures are grounded on pedagogical criteria, besides more usual graph matching.

Regarding such measures: (1) *Overlapping Edges Degree* considers nodes shared between CMs and linked by identical relationships; (2) *Prerequisite Constraints* computes similarity by analyzing common learning paths; (3) *Knowledge Commonality* evaluates the *centrality* that a given concept possesses in the compared maps.

In their work [28], the authors introduce a novel approach to assessing didactic similarity among various CM concepts, leveraging DL. Utilizing the DL InferSent algorithm, which relies on sentence embedding, they infer similarity between concepts and maps, taking into account the concepts only.

In [29], the authors propose a technique for matching elements or subsets of CMs, employing a similarity flooding algorithm to facilitate the comparison and fusion of CMs, aligning nodes and substructures across simulated CMs and CMs created by students.

In contrast to existing approaches in the literature, our method for measuring similarity utilizes embedding techniques [10,11,30] to assess the semantic resemblance between two CMs by considering each relationship's basic elements (the relation and the concepts involved in it), as well as a holistic perspective.

### *2.4. Support to Learning (and Teaching) Analytics*

One of the aims of this paper is to present an approach to the process of teaching and learning (1) based on the CM similarity measure we proposed in this paper, (2) inspired, from the theoretical viewpoint, by the advancements in Learning Analytics (LA) and Teaching Analytics (TA) [31–33], and (3) using, from the practical viewpoint, the dashboard technology [13,14].

LA, and in particular TA, have been fields of study for several years, yet they are still open to systematization. TA is a sub-field of LA, focusing on the design, development, and evaluation of visual analytics methods and tools to support teachers in their inquiry activity. In effect, LA, as a research and application field, is quite interdisciplinary, being linked and applying methodologies and tools coming from other research areas (such as Education, Pedagogy, Human–Computer Interaction, User Modeling, Recommender Systems, Machine Learning, and Artificial Intelligence).

In [31], a *reference model* is proposed as a guide in the development of LA initiatives. The model advocates the use of a four-dimensional classification schema: *'data and environments (what?), stakeholders (who?), objectives (why?), and methods (how?)'*. The results of a literature review based on such a model allowed the conclusion that, among other aspects, LA solutions receive data mainly from Intelligent Tutoring Systems and Learning Management Systems by means of classification and prediction tools.

In [32], a *consolidated model* for LA is presented, where three dimensions of interconnected and dependent features concur to form a rational and well-founded solution for LA. The first dimension, *Theory*, deals with the necessity to found all the data-drive activities of an LA system on a consistent and consistently applied theory. In other words, the *'Potential to analyze and identify patterns in large datasets'* must be guided by a theory.

As a source for such theoretical foundation in an LA system, Psychology, Sociology, Organization Science, Information Science, and Semiotics are pointed out. Also, an application in the educational field could help advance the theory.

The second dimension is the *Design* one, which is in turn discussed as a tri-parted area: First, there is a design activity regarding *Interaction and visualization* to be fulfilled in the LA system regarding the support for the users (learners and teachers mainly) for the actual acquisition of insights out of an efficient interface. Second, the LA system must also be founded on a design of the *Learning* components, which should support effective learning by allowing the general system to select, extract, analyze, and visualize data about the effects of the learning activities, which will be useful to give insights only if this process was theoretically sound, and so giving valid results. Third, a *Study* design is needed that should allow data research and evaluation during and after the use of the system (also for longitudinal studies, for instance). In the authors' words: *'A study design consideration should understand the nature of data collection, possible ways that the data can be analyzed, and the types of questions to be answered'*.

The review in [34] discusses the evidence available in the literature regarding the results of using LA systems on the learner's success. In particular, the paper shows results about how successful study continuation and completion have been enhanced and how dropouts have been influenced for the better by the application of LA. However, it seems that the available data are usually related to ad hoc experiments, while a large-scale investigation proposing methods to adapt the LA offer to the particular characteristics and possibilities of the school or university is still lacking.

In conclusion of this section, regarding research related to this article's subject, we may summarize the main points we intend to provide with this article as follows:

- Regarding the use of a similarity measure to compare CMs, to the best of our knowledge, no approach is provided in the literature where concepts and relationships of a CM are managed through the DL techniques we proposed. Our DL-based approach can surely be further deepened and elaborated, so we also hope to have provided some means to further deepen this type of approach.
- While the use of similarity evaluation between maps is considered, it appears always bound to verify the correctness of the maps and/or establish a classification of the maps on that basis. Such evaluation can be performed either manually (through the application of a rubric) or by automatic assessment, and similarity considerations are usually done against a model map (teacher's map or maps) rather than among the students' maps.

The support for a direct (and similarity-directed) analysis of the students' maps, aiming to let the teacher obtain a comprehensive idea of how the students are distributing their efforts within the class, is altogether missing in the literature.

We believe that, despite being disconnected from direct map-correctness assessment, such a feature can support the teacher's analysis of class trends: showing how possible clusters of similar CMs are present in the students' products can allow the observation of how the domain knowledge is differently organized in the class, and to appreciate students' common interpretations (however right or wrong they might be). This might be done, for instance, by visiting sample maps in a group of similar ones, providing the teacher with a feasible way to analyze the whole set of the class' CMs.

## 3. Computation of Map Similarity through Deep-Learning Methods

Sentence-embedding models aim to capture the semantic core of a sentence in a fixed-length vector, offering a significant advancement over traditional text representation methods like Bag-of-Words (BoW) or one-hot encoding. Unlike these older techniques, sentence embeddings can understand the context, meaning, and relationships among words, providing a more nuanced and effective means of representing text data for various natural-language processing tasks. For this reason, sentence embedding is a well-studied area with dozens of proposed methods [35]. Of all the models proposed, we chose two: *InferSent* and *Universal Sentence Encoder* (USE). The former was chosen because it has already been used with good results in other works on CMs (see, for example, [28]). The latter was chosen because studies have shown its validity in different research fields in the treatment of semantic similarity between sentences [36].

We used both approaches to construct each CM embedding, allowing us to compare them and assess the most effective method for computing CM similarity. Each CM was divided into individual sentences, and then each of them was transformed into embedding. In this way, each CM was represented as a set of vectors of real numbers with size 512.

### 3.1. InferSent Sentence-Embedding Encoder

InferSent is a technique for encoding sentence embeddings that offers semantic representations of English sentences, trained on natural-language inference data, and that generalizes well to many different tasks [10].

This is a supervised model based on DL, trained with the Stanford Natural-Language Inference (SNLI) corpus. The SNLI corpus consists of 570,000 English sentence pairs labeled as implying, contradicting, or being neutral to each other. InferSent has two versions: the original uses *GloVe* vectors for pre-training word embeddings (https://nlp.stanford.edu/projects/glove/ accessed on 23 October 2023) [30], while the updated version adopts *fastText* for the same goal. In our research, we employed fastText, a tool created by Facebook's AI Research (FAIR) lab for developing word embeddings and text classification (https://fasttext.cc/docs/en/crawl-vectors.html accessed on 31 August 2023). FastText aims to encode full sentences to maintain their semantic integrity. As the development environment, the model was implemented in Python language, using the CoLab tool provided by Google (https://colab.research.google.com/ accessed on 15 September 2023).

### 3.2. Universal Sentences Encoder Embedding

The *Universal Sentence Encoder* (USE) transforms text into high-dimensional vectors, making them applicable for tasks such as text classification, semantic similarity evaluation, and other natural-language processing (NLP) activities [11]. This pre-trained DL model takes English text of variable length and produces a 512-dimensional vector of real numbers. USE incorporates the *Deep Averaging Network* (DAN) encoder, supported by the TensorFlow Python library. The DAN model, based on feed-forward neural network principles, employs either Word2Vec or GloVe for its embedding method. It calculates the sum or average of embeddings for each word in a sentence and then channels these through one or more layers of feed-forward networks, specifically Multi-Layer Perceptrons (MLPs). The final classification is executed in the last layer using logistic regression, Naïve Bayes, or Support Vector Machine (SVM) techniques. In this case, as the development environment, the model was implemented in Python using the CoLab tool provided by Google.

### 3.3. Concept-Map Embedding

As discussed in Section 2, comparing two CMs is a complex task with various challenges, partly dependent on the perspective from which similarity is evaluated. The CM, depicted as a directed graph, requires an algorithm for measuring its difference from another CM, which must be based on one or more criteria. Each criterion emphasizes a preference for a specific aspect of similarity.

To encode a CM using embeddings, we initially consider that each edge $r$ in the map, linking two nodes $h$ and $t$ (with $r$ directed from $h$, the head concept, to $t$, the target concept), can be encapsulated as a Relationship Triplet $RT = (h, r, t)$. Therefore, a CM, denoted as $C$, can be described as the collection of its relationship triplets. Figure 2 shows the process, where both the embedding models are used for representing each CM.

$$C^T \equiv \{(c_h, r_{c_h,c_t}, c_t), \forall edge r_{c_h,c_t} \in C\}$$

where $r_{c_h,c_t}$ is the edge in $C$ that connects the concepts $c_h, c_t$. (We can quite reasonably use the condition of non-existence for isolated concepts in a map). Then, we express each relationship triplet in $C^T$ as a *sentence*: let be the sentence defined by the relational triplet $(h, r, t)$:

$$s^C_{h,r,c} =< c_h > + < r > + < c_t >$$

If $n_C$ is the number of sentences so obtained, $s_i, i \in [1 \cdots n_C]$ the map $C$ can then be represented as the set of sentences $C^S = \{s_1^C, s_2^C, \ldots, s_{n_C}^C\}$. Table 1 shows the $S^{Molecules}$ set (for the map in Figure 1), also giving an idea of the sentence's appearance.

Since each sentence can be represented as an embedding (see just a few lines later), $C$ is then straightforwardly represented as a set of embeddings:

$$C^{Emb} = \{embedding(s_1^C), embedding(s_2^C), \ldots, embedding(s_{n_C}^C)\}$$

**Table 1.** The $S$ set for the $CM$: *Molecules*.

| ID | Sentence |
|---|---|
| $s_1$ | molecules can be water |
| $s_2$ | water found in living things |
| $s_3$ | molecules have motion |
| $s_4$ | motion determines states |
| $s_5$ | motion increased by heat |
| $s_6$ | states can be solid |
| $s_7$ | states can be gas |
| $s_8$ | states can be liquid |
| $s_9$ | living things can be plants |
| $s_{10}$ | living things can be animals |
| $s_{11}$ | water can change states |

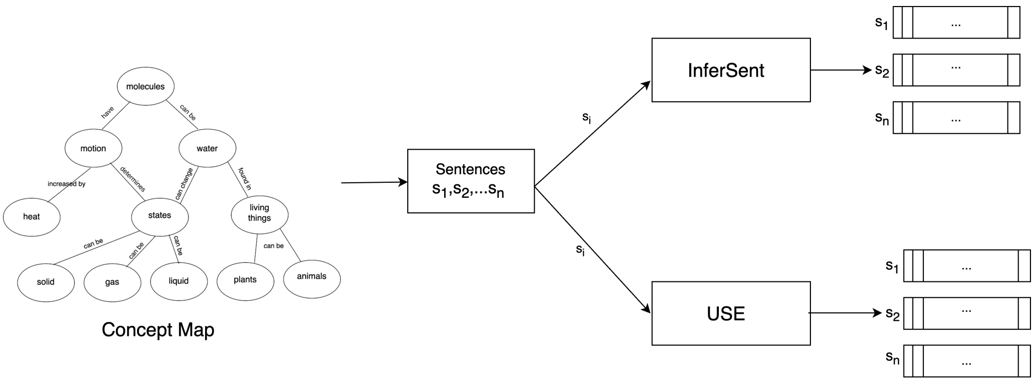


**Figure 2.** The CM embedding process: first, each CM is transformed into a set of sentences $S$. Second, each sentence $s_i$ is transformed into its embedding by both InferSent and USE.

According to the research aims expressed in Section 1, to deal with the similarity of two CMs, $C1, C2$, here we aim to use the aforementioned embedding techniques, InferSent and Universal, to achieve the following goals:

1. Producing the embedding representations of $C1$ and $C2$, respectively $C1^{Emb}$ = {embedding($s_1^{C1}$), embedding($s_2^{C1}$), …, embedding($s_{n_{C1}}^{C1}$)} and $C2^{Emb}$ = {embedding($s_1^{C2}$), embedding($s_2^{C2}$), …, embedding($s_{n_{C2}}^{C2}$)}.
2. Computing a matrix:
$$EMB(C1, C2) = \{d_{ij}\}$$
where $d_{ij} = distance(s_i^{C1}, s_j^{C2})$, for $i \in [1 \cdots n_{C1}], j \in [1 \cdots n_{C2}]$ (each $d_{ij}$ is the distance between the $i$-th sentence of $C1$ and the $j$-th sentence of $C2$).
Notice that in the experimental evaluation, we used both the Euclidean and Cosine distance.

Notice also that the distances are computed on the embedding representation of the sentences.

3. Computing the arithmetic mean and the Standard Deviation of the distances of all the sentences in $C1$ from all those in $C2$ (distances computed on the embeddings): d($C1$, $C2$) = Mean($\{$v, v$\in EMB(C1, C2)\}$), $STD(C1, C2) = Stdev(\{$v, v$\in EMB(C1, C2)\})$.
4. Implementing a system, based on our similarity measure, to provide teachers with a dashboard to monitor concept-mapping activity.

In fact, d($C1$, $C2$) is the measure we propose for the similarity between the two CMs.

In the following, we use both the InferSent and Universal DL embedding models as the embedding methods for sentences to compare their use and give indications about their usefulness. Moreover, as the *distance()* function mentioned in point 2, we used both the Euclidean and Cosine distances separately, again to compare them during the evaluation of our experimental data.

## 4. Experimental Evaluation

We conducted an experimental activity to evaluate how the sentence-embedding approaches, described in the previous section, can measure the similarity of two CMs and, along the way, assess which one of the two approaches is better performing out of the one using InferSent and the one applying USE. To do this, we verified how the measures aligned with the assessments from a sample of teachers. Assuming the teacher's assessment of similarity as the baseline, the comparison with the automated assessments provided by the measures would then allow the answering of our RQ.

We selected a sample of teachers from whom we asked for several map-similarity assessments. The teachers, selected randomly from a group of primary and secondary-school grade educators, were 31.

To compose a questionnaire, where several map comparisons would be evaluated by the sample, we also selected a set of CMs retrieved from the World Wide Web. Eventually, 5 maps were included in this set: *Living Things*, *Molecules*, *Water*, *Neural Networks*, and *Plant*. The maps are not presented here for reasons of space; however, they are visible in the questionnaire, which is referenced in Section 4.4.

This allowed the proposal of the sample with 10 pairs of maps to be compared. The general criterion for map selection was to have maps granting a variety of similarity results and being neither too small nor too big (such as between 10 and 20 nodes). In addition, the number of selected maps was stated to limit the number of comparisons and, accordingly, the overall workload imposed upon the respondents.

Figure 3 illustrates the experimental procedure: basically, the same comparisons and assessments of similarity were to be conducted using the sentence-embedding approaches (four times, overall, applying InferSent and USE, each one by two definitions of distance) and once by the sample. Then, each one of the four sentence-embedding-based datasets would be compared with the teachers' one.

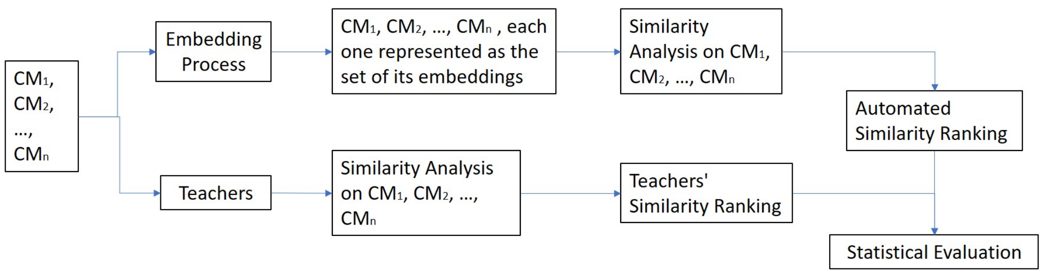


**Figure 3.** The workflow of the experimental procedure.

### 4.1. Environment Setting

To implement, in Python programming language, the sentence-embedding approaches, we employed the *CoLab* service provided by Google (https://colab.research.google.com/

accessed on 20 August 2023). This free, web-based service allows the development of computer programs using Python software, 3.1 libraries for DL, basically through a web browser, with very limited needs for configuration and access to powerful computational tools. It is a "Notebook"-based service: it relies on the implementation of a system through logical software blocks (the notebooks) developed according to the *Jupyter* software model (https://jupyter.org/ accessed on 17 August 2023).

We adopted this development service to enhance the availability of computational resources that we would not have locally. In particular, we had the chance, through Google CoLab, to use standard Python libraries (such as numpy and tensorflow), also exploiting web-accessible GPU and parallelization resources to make our embedding models run.

*4.2. Setting of the Experimental Data*

Since we wish to compare the assessment of similarity produced through our measure with the one provided by the teachers' sample, we organized two phases of data production. In the first phase, we produce the similarity assessments, according to our measure, for each one of the map comparisons proposed in the questionnaire presented to the sample. The second phase is the collection of teachers' assessments on the same map comparisons.

During the first phase:

1. for each one of the 5 experimental maps, say $C$, we build the corresponding sentence representation $C^S = \{s_1^C, s_2^C, \ldots, s_{n_C}^C\}$ (cfr. Section 3). This is carried out using a parsing Python module that runs in Colab.
2. then, the embedding representation is computed as: $C^{Emb}$ = {embedding($s_1^C$), embedding($s_2^C$), ..., embedding($s_{n_C}^C$)}, as explained in the following subsection.
   Since we intended to use two embedding models (USE, and InferSent), we would compute two embedding representations for each one of the CMs, accordingly.
3. then, based on our similarity measure, the distances between each pair of maps involved in the teacher's sample comparison tasks are computed. This computation is performed using two metrics for the distance between vectors (the embeddings): the Cosine and the Euclidean distance.
   Therefore, eventually, for each comparison request for the teacher's sample, we have four similarity assessments (coming from the use of two embedding representations with two different vector distance metrics).

*4.3. Computation of the Embeddings and the Distances*

To produce the evaluation data, we used both USE and InferSent, applying the pre-trained model with no fine-tuning. In both cases, we represented a sentence in a CM with a 512-feature vector (the embedding).

Considering one of the two CMs to compare, $C$, the embedding representation is: $C^{Emb}$ = {embedding($s_1^C$), embedding($s_2^C$), ..., embedding($s_{n_C}^C$)} where embedding($s_i^C$) represents the i-th sentence in $C^S$ (C's sentence representation).

Then, we computed the distance matrix EMB($C1$,$C2$) = $\{d_{ij}\}$ for each pair ($C1$, $C2$) of CMs involved in the experiment, and determined the related similarity value as the arithmetic mean of the distances of all the sentences in $C1$ from all those of $C2$ (distance computed on the embeddings): d($C1$, $C2$) = Mean($\{$v, v$\in EMB(C1, C2)\}$).

As mentioned earlier, we have four sets of data here, as we computed the Euclidean distance as well as the Cosine distance in both the InferSent and USE embedding computations.

Table 2 presents the Cosine distances among the sample maps when the embeddings are computed through the USE model; similarly, Table 3 is about the Euclidean distances, still when USE is applied. Taking into consideration the fact that CMs are usually limited to a relatively small number of concepts, and relationships, the cost of computation for this step of our process is not significantly high.

**Table 2.** The Cosine distances among all CMs using the Universal DL model.

| vs. | Living | Molecules | Water | Neural | Plant |
|---|---|---|---|---|---|
| living | - | 0.3846 | 0.2829 | 0.1136 | 0.2042 |
| molecules | - | - | 0.2879 | 0.1202 | 0.1838 |
| water | - | - | - | 0.0869 | 0.1678 |
| neural | - | - | - | - | 0.1118 |
| plant | - | - | - | - | - |

**Table 3.** The Euclidean distances among all CMs using the Universal DL model.

| vs. | Living | Molecules | Water | Neural | Plant |
|---|---|---|---|---|---|
| living | - | 1.0654 | 1.1693 | 1.3297 | 1.2586 |
| molecules | - | - | 1.1662 | 1.3246 | 1.2751 |
| water | - | - | - | 1.3495 | 1.2876 |
| neural | - | - | - | - | 1.3314 |
| plant | - | - | - | - | - |

### *4.4. Teachers' Assessment on the Experimental Comparisons*

A sample of teachers was asked to complete an online questionnaire using the Google Module platform (Available at: https://forms.gle/QvappMpWXLiwYuL28 accessed on 15 May 2023). The survey included ten questions, each requiring the comparison of two CMs. These CMs were displayed, and teachers were asked to evaluate their similarity on a 5-point Likert scale. We prepared ten comparisons using five different CMs. The Likert scale ranged from *very different* to *very similar* at the extremes. Teachers were advised to use specific criteria to guide their evaluations (refer to Figure 4). The degree of similarity between the two CMs was assessed based on the fact that:

1. they have concepts in common (and in this case also, synonyms of concepts are to be taken into consideration);
2. they present common semantic links between common concepts (and possible synonymy between semantic relations is to be taken into account).

In Figure 5, a screenshot of the questionnaire results is shown.

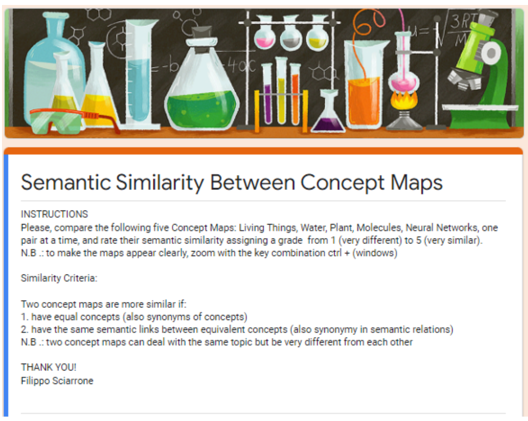


**Figure 4.** The home page, showing instructions, of the questionnaire submitted to the sample of teachers.

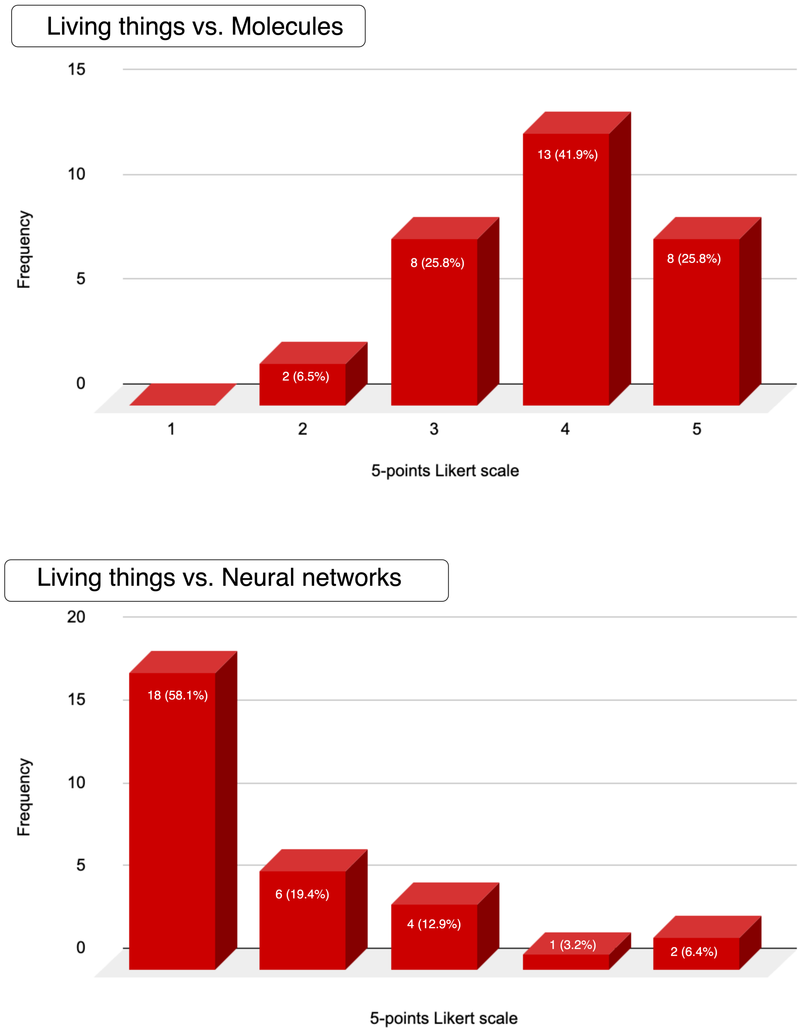


**Figure 5.** The results of the questionnaire for *Living things vs. Molecules* and *Living things vs. Neural networks* comparisons.

### *4.5. Statistical Evaluation*

Based on the questionnaires we received, we ranked all similarity distances. This resulted in three distinct rankings for all pairs of CMs: one based on Cosine distance, one on Euclidean distance, and one derived from teachers' assessments. To examine the relationships or lack thereof between the distances suggested by teachers versus those proposed by two DL models, we computed the non-parametric *Spearman's rank correlation coefficient* [37]. The evaluations by teachers served as the baseline for comparing the rankings produced by the DL algorithms. Table 4 displays the correlation outcomes. There is a strong correlation between the scores from the Universal model and the teachers' assessments. However, the correlation between the teachers' assessments and the conclusions drawn from the InferSent model is weaker. Additionally, we conducted a one-tail test, which yielded a of $p$-value = 0.004 for the Universal model when compared with both the Cosine and Euclidean distances.

**Table 4.** Spearman's correlation results.

| | InferSent (Cosine) | InferSent (Euclidean) | Universal (Cosine) | Universal (Euclidean) |
|---|---|---|---|---|
| $\rho$ | 0.51 | 0.46 | 0.85 | 0.85 |

## 5. The Teacher Dashboard

After the successful verification of the RQ, here we move on to show the concept of a dashboard application built on the sentence-embedding-based method, which is designed for the teacher, with the aim to support the preliminary analysis of the CMs produced by the class. The nature of such preliminary analysis is described in the final part of Section 1. As stated in Section 1, the dashboard we present is not intended as (part of) a fully fledged system ready to be applied in the class environment, rather it should provide the reader with a clear suggestion of how it can be useful for the teacher, as an application of the similarity measure we evaluated, during the task of preliminary analysis of the (possibly many) maps submitted by the class. Hence, the dashboard is intended to be integrated into a whole process of support to Learning by concept mapping and Teaching Analytics, which can be represented by the six-step cyclic process shown in Figure 6 [33].

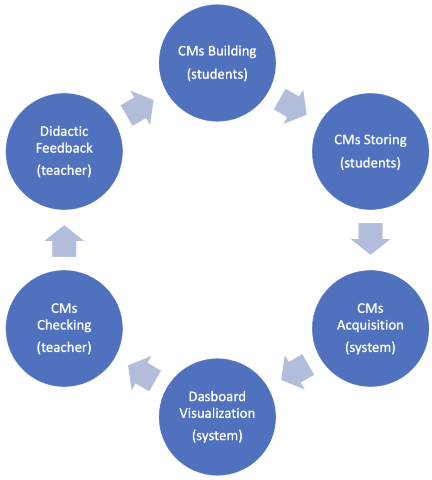


**Figure 6.** The six-step process to improve Learning and Teaching Analytics for concept-mapping activity.

As the figure shows, the process is based on the following six steps, forming a didactic cycle to improve the concept-mapping activity:

1. *CMs Building*. First, the students build their own CM in a graphic format and second in text format. Text format means a text file where each row contains the textual representation of the sentence of the CM, as shown in Table 1. This format will allow a CM to be processed by the InferSent and USE DL models to produce the embeddings and then compute the distances between CMs, as explained below.
2. *CMs Storing*. The students can deliver their files through a common area, such as a folder on Google Drive.
3. *System Acquisition*. The system automatically takes the CMs as input both in text and graphic format. In this phase, the neural engine processes all the text files representing the CMs delivered by the students. This process is shown in Figure 7.
   Let $B = \{b_1, b_2, \ldots, b_m\}$, the set of sentence embeddings for the teacher's CM (that is our baseline), and $E = \{e_1, e_2, \ldots, e_n\}$ the set of generic embeddings sentences for a given student's CM. Hence, each learner's CM can be represented by a single value as $\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} d(b_i,e_j)}{m*n}$, where $d(b_i, e_j)$ is the Euclidean distance computed between the teacher's map's set of embeddings and the learner's one, and this value represents the global distance between the teacher's and learner's CMs. Notice that in this process, we are using only the Euclidean distance to compute embedding distances, as (cf. previous section) Cosine and Euclidean distances have been revealed to have similar behaviors.
4. *Dashboard Visualization*. The system draws a scatterplot, where each point represents a learner's $CM$, by the coordinates $(x, y)$ where:

- the coordinate $x$, on the horizontal axis, is the order number of the student (map) in the class,
- the coordinate $y$, on the vertical axis, is the distance of the corresponding student's map from the teacher's CM.

In other words, the farther the point is from the X-axis, the lower the similarity of the learner's CM with the teacher's.

5. *CMs Checking*. The teacher, by clicking on a particular point, can visualize the corresponding $CM$. This process is shown in Figure 8, where a prototypical system is proposed. In the example proposed by the figure, the baseline CM was *Trees*. The teacher wrote her own CM while her students (nine in this case) built their own. By observing the scatterplot, the teacher wants to check the *farthest* $CM$. Then, by clicking on the red point, the system opens the relative $CM$.
6. *Didactic Feedback*. After the analysis performed in the previous step, the teacher will be able to provide the learner (author of the map) with the necessary feedback.

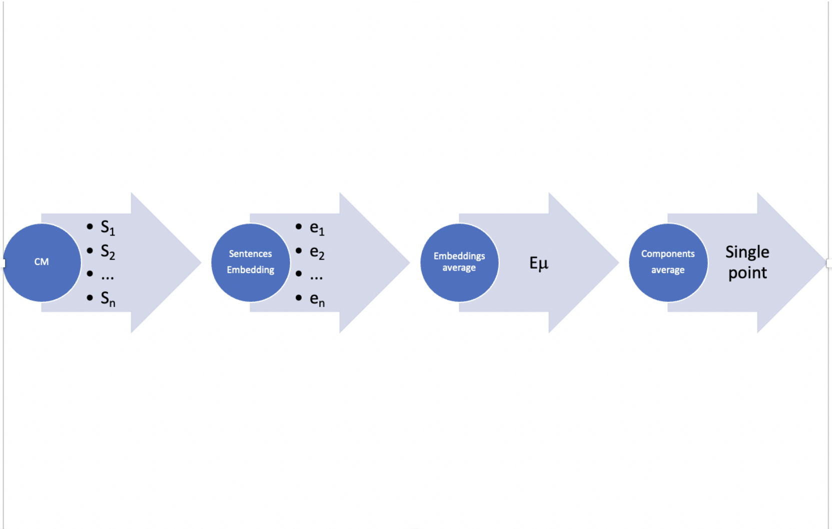


**Figure 7.** The CM reduction to a single-point process.

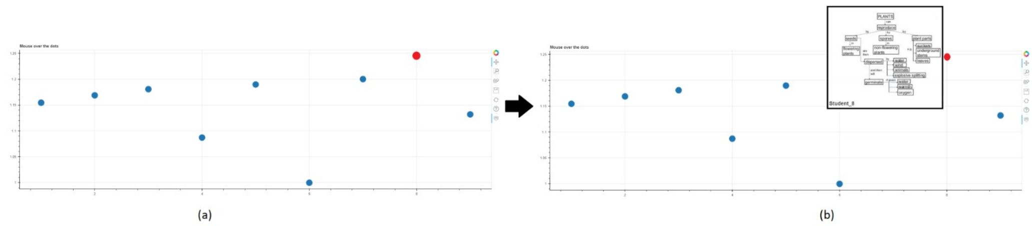


**Figure 8.** The teacher dashboard as a scatterplot where each blue point represents a student's CM. Instead, the farthest CM from the teacher's one is represented as a red point in (**a**). By clicking the red point, one can check on the student's CM and give the appropriate feedback to improve the CM building activity (**b**).

## 6. Discussion

Here, we elaborate on the experimental results of the similarity measure, trying also to make some conclusive points about the usefulness of the dashboard concept we presented.

In Section 2, we expounded on the research efforts present in the literature and related to this work. As noted in that section, there seems to be a lack of studies on the use of DL-based methods to compute measures of similarity (or distances) for CMs, so we have actually limited possibilities to compare our proposed measure with other measures. For our measure in this article, our plan was to validate it against a sample of teachers, which provided the baseline. Adding a comparison with the results of other measures, possibly based on the same baseline, would be a further step in our work, although it would also be complex, as we should secure implementations of such other measures. For the above-mentioned reasons, we discuss only the results from the experimentation and present no further comparisons.

### 6.1. Research Question

The RQ basically asks whether a similarity measure for CMs can be defined through DL embedding models to satisfy the pedagogical viewpoint of the teacher. The measure of similarity we presented in Section 3 intended to capture the structure and semantics of a CM using methods to produce embeddings as a representation of the map and distance measures between embeddings. We used two embedding methods and two distance measures between embeddings to make comparisons and see the validity of the general approach.

We selected a sample of five CMs, from which we generated ten comparisons. These were presented to a group of 31 teachers, who were tasked with evaluating the similarity of the CMs across the ten comparisons.

The experimental results, summarized in Table 4, indicate a strong correlation between the similarity scores computed using the Universal DL model and those provided by the sample of teachers. The correlation holds for both the Cosine and Euclidean distance, which we used to compute the distance between embedding sets. Thus, the methodology implemented, which relies on embeddings, is capable of capturing the semantics of the sample of CMs. Consequently, it effectively supports teachers by providing similarity assessments that closely align with those made by human teachers in the sample.

A further insight is derived from the $p$-value discussed at the end of the previous subsection. We posited the null hypothesis, $H_0$, that there is no monotonic correlation between the two statistical distributions: one defined by teacher assessments and the other by automatic computations. Conversely, the alternative hypothesis, $H_1$, proposed the existence of a monotonic correlation between these distributions. Given that $H_0$ was rejected (with a $p$-value of 0.004), we can regard the experiment as an effective, albeit preliminary, validation of the embedding-based approach to measuring similarity.

### 6.2. Teacher's Dashboard Expected Usefulness

Of course, the measure we proposed cannot help the teacher without being integrated into an overall system, such as a fully fledged application realizing the process described in Section 5. To give some elements that appreciate the potential usefulness of using our DL-based measure, we described the design of a dashboard that could be embedded in the above-mentioned system. Therefore, the dashboard is a proof-of-concept implementation providing a visual rendering of the many CMs produced by a class of learners as points in the bi-dimensional space, according to the distances computed among them through the similarity measure.

The integration of the dashboard mechanism into a complete Learning support system is a matter for future work; however, we received indications from the teachers participating in the experimental evaluation that we presented about what feature would be more attractive, to be included in the implementation. Figure 8 shows a case study with its usefulness: by this system, the teacher can check each CM and give immediate feedback to the student. The system is planned to be extended in a client-server architecture to automate the CM delivery phase through a suitable interactive interface.

## 7. Conclusions and Future Work

In this paper, we investigated the potential of using DL techniques to measure the similarity between CMs. This approach addresses challenges associated with various perspectives on the concept of similarity, such as comparing graphs. We specifically utilized InferSent and the Universal Sentence Encoder, both of which are well-suited for generating embeddings from incoming sentences. Initial experiments have shown a positive correlation between these techniques and the assessments made by a sample of teachers. Notably, the Universal Sentence Encoder closely aligns with the teachers' judgments. Looking ahead, we plan to conduct more extensive experiments and compare these methods with other embedding techniques.

In terms of the possible impact of the similarity measure we proposed, we might add some notes about possible different applications. In general, any system where the semantic collection of texts is managed through the use of DL could have a use for similarity measure as the one we propose: for instance, machine-learning and DL techniques are used to carry out sentiment analysis, and news/opinions check and comparison (see, for instance, [38]). Other interesting fields remaining in a Technology Enhanced Learning setting could be (1) the Automated Grading of a CM by analysis of its correctness (which was not considered in this article) and its similarity to a teacher's CM; (2) representation of a course's learning corpus in a Large Language Model to support conversational interaction between student and a chatbot; and (3) support the semi-automated evaluation of essays, again by embedding representations of the students' essays and the teacher's one (or ones).

Apart from such relatively distant applications, we think that the research we presented here should be followed up, further elaborating on its usefulness, definition details, and applicability. We can point out that a task of interest is in comparison with other methods of similarity analysis for CMs, although they are likely to be not based on DL techniques. This must be done by building a new experimental setting, still involving teachers to create a baseline, and then carefully obtaining the results from our and others' measures to make a comparison and state some conclusions about the efficiency and correctness of the computations.

A second line of work would be, of course, the implementation of a system where the measure of similarity would be actually used to support the whole learning and teaching cycle described in Section 5. That would allow the extension of our experimental work to the direct application of classes with students and teachers.

Additional important lines of inquiry will regard two topics in particular:

1. Once a system is actually available, analysis of its usability and effectiveness on teachers and learners could be pursued.
2. Since we build the embeddings based on simple sentences extracted from the map, involving 2 concepts and one relationship, it is conceivable that we might reach a better semantic representation of the map, involving embeddings on longer sequences (such as whole "branches" of concepts and relationships). This is an endeavor to be carefully planned, in particular with an eye on the possible computational impact consequent to the production of longer sequence embeddings.

**Author Contributions:** Conceptualization, F.S. and M.T.; Methodology, F.S.; Validation, M.T.; Formal analysis, M.T.; Investigation, F.S.; Writing—review & editing, M.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially funded by Universitas Mercatorum grant number 16-FIN/RIC.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Albino, V.; Berardi, U.; Dangelico, R. Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *J. Urban Technol.* **2015**, *22*, 3–21. [CrossRef]
2. Gomede, E.; Gaffo, F.; Briganó, G.; De Barros, R.; De Souza Mendes, L. Application of Computational Intelligence to Improve Education in Smart Cities. *Sensors* **2018**, *18*, 267. [CrossRef] [PubMed]
3. Liu, D.; Huang, R.; Wosinski, M. 1—Development of Smart Cities: Educational Perspective. In *Smart Learning in Smart Cities*; Lecture Notes in Educational Technology; Springer: Berlin/Heidelberg, Germany, 2017.
4. De Marsico, M.; Sciarrone, F.; Sterbini, A.; Temperini, M. Supporting mediated peer-evaluation to grade answers to open-ended questions. *EURASIA J. Math. Sci. Technol. Educ.* **2017**, *13*, 1085–1106.

5. Limongelli, C.; Sciarrone, F.; Temperini, M.; Vaste, G. LECOMPS5: A web-based learning system for course personalization and adaptation. In Proceedings of the IADIS International Conference e-Learning, Amsterdam, The Netherlands, 22–25 July 2008; pp. 325–332.
6. de Ries, K.; Schaap, H.; van Loon, A.M.; Kral, M.; Meijer, P. A literature review of open-ended concept maps as a research instrument to study knowledge and learning. *Qual. Quant.* **2022**, *56*, 73–107. [CrossRef]
7. Buhmann, S.; Kingsbury, M. A standardized holistic framework for concept map analysis combining topological attributes and global morphologies. *Knowl. Manag. E-Learn.* **2015**, *7*, 20–35.
8. Limongelli, C.; Marani, A.; Sciarrone, F.; Temperini, M. Measuring the Similarity of Concept Maps According to Pedagogical Criteria. *IEEE Access* **2022**, *10*, 27655–27669. [CrossRef]
9. Marriott, R.; Torres, P. Formative and Summative Assessment of Concept Maps. In *Innovating with Concept Mapping. CMC 2016*; Cañas, A., Reiska, P., Novak, J., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2016. [CrossRef]
10. Conneau, A.; Kiela, D.; Schwenk, H.; Barrault, L.; Bordes, A. Supervised Learning of Universal Sentence Representations from Natural Language Inference Data. *arXiv* **2017**, arXiv:1705.02364.
11. Cer, D.; Yang, Y.; Kong, S.; Hua, N.; Limtiaco, N.; John, R.S.; Constant, N.; Guajardo-Cespedes, M.; Yuan, S.; Tar, C.; et al. Universal Sentence Encoder. *arXiv* **2018**, arXiv:1803.11175.
12. Montanaro, A.; Sciarrone, F.; Temperini, M. A Deep Learning Approach to Concept Maps Similarity. In Proceedings of the 2022 26th International Conference Information Visualisation (IV), Vienna, Austria, 19–22 July 2022; pp. 1–10.
13. Park, Y.; Jo, I. Development of the Learning Analytics Dashboard to Support Students' Learning Performance. *J. Univers. Comput. Sci.* **2015**, *21*, 110–133.
14. Susnjak, T.; Ramaswami, G.S.; Mathrani, A. Learning analytics dashboard: A tool for providing actionable insights to learners. *Int. J. Educ. Technol High Educ.* **2022**, *19*, 12. [CrossRef]
15. Novak, J. The importance of conceptual schemes for teaching science. *Sci. Teach.* **1964**, *31*, 10–13.
16. Ausubel, D.; Novak, J.; Hanesian, H. *Educational Psychology: A Cognitive View*, 2nd ed.; Holt, Rinehart and Winston: Austin, TX, USA, 1978.
17. Bruner, J. Child's talk: Learning to use language. *Child Lang. Teach. Ther.* **1983**, *1*, 111–114. [CrossRef]
18. Vygotsky, L.S. *Thought and Language*; MIT Press: Cambridge, MA, USA, 1962.
19. Ausubel, D. *The Psychology of Meaningful Verbal Learning*; Grane and Stratton: New York, NY, USA, 1963.
20. Ionas, I.G.; Geana, M.V. graphed: A Web-Based Concept Mapping Application for Instruction and Research. In Proceedings of the Seventh International Conference on Higher Education Advances, València, Spain, 22–23 June 2021; pp. 563–570.
21. Cañas, A.J.; Valerio, A.; Lalinde-Pulido, J.; Carvalho, M.; Arguedas, M. Using WordNet for word sense disambiguation to support concept map construction. In Proceedings of the String Processing and Information Retrieval, Manaus, Brazil, 8–10 October 2003; Springer: Berlin/Heidelberg, Germany, 2003; pp. 350–359.
22. Cañas, A.; Bunch, L.; Novak, J.; Reiska, P. CmapAnalisys: An Extensible Concept Map Analysis Tool. *J. Educ. Teach. Trainers* **2013**, *4*, 36–46.
23. Stevenson, M.; Hartmeyer, R.; Bentsen, P. Systematically reviewing the potential of concept mapping technologies to promote self-regulated learning in primary and secondary science education. *Educ. Res. Rev.* **2017**, *21*, 1–16. [CrossRef]
24. Hay, D.; Kinchin, I. Using concept maps to reveal conceptual typologies. *Educ. Train.* **2006**, *48*, 127–142. [CrossRef]
25. Kinchin, I.; Cabot, B. Reconsidering the dimensions of expertise: From linear stages towards dual processing. *Lond. Rev. Educ.* **2010**, *8*, 153–166. [CrossRef]
26. Novak, J.D. *Learning How to Learn*; Cambridge University Press: Cambridge, UK, 1984.
27. Sasson, E.; Ravid, G.; Pliskin, N. Improving similarity measures of relatedness proximity: Toward augmented concept maps. *J. Inf.* **2015**, *9*, 618–628. [CrossRef]
28. Limongelli, C.; Schicchi, D.; Taibi, D. Enriching Didactic Similarity Measures of Concept Maps by a Deep Learning Based Approach. In Proceedings of the 2021 25th International Conference Information Visualisation (IV), Sydney, Australia, 5–9 July 2021; pp. 261–266. [CrossRef]
29. Marshall, B.; Chen, H.; Madhusudan, T. Matching knowledge elements in concept maps using a similarity flooding algorithm. *Decis. Support Syst.* **2006**, *42*, 1290–1306. [CrossRef]
30. Pennington, J.; Socher, R.; Manning, C.D. GloVe: Global Vectors for Word Representation. In Proceedings of the Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
31. Chatti, M.; Dychkhoff, A.; Schroeder, U.; Thus, H. A reference model for learning analytics. *Int. J. Technol. Enhanc. Learn.* **2012**, *4*, 318–331. [CrossRef]
32. Gasevic, D.; Kovanovic, V.; Joksimovic, S. Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learn. Res. Pract.* **2017**, *3*, 63–78. [CrossRef]
33. Sergis, S.; Sampson, D. Teaching and Learning Analytics to Support Teacher Inquiry: A Systematic Literature Review. In *Learning Analytics: Fundaments, Applications, and Trends*; Studies in Systems, Decision and Control; Peña-Ayala, A., Ed.; Springer: Cham, Switzerland, 2017; Volume 98. [CrossRef]
34. Ifenthaler, D.; Yau, J. Utilising learning analytics to support study success in higher education: A systematic review. *Educ. Tech Res. Dev.* **2020**, *68*, 1961–1990. [CrossRef]

35. Mistry, D.M.; Minai, A.A. A Comparative Study of Sentence Embedding Models for Assessing Semantic Variation. *arXiv* **2023**, arXiv:2308.04625.
36. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv* **2019**, arXiv:1908.10084.
37. Siegel, S.; Castellan, N. *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed.; McGraw–Hill, Inc.: New York, NY, USA, 1988.
38. Poornima, A.; Priya, K.S. A Comparative Sentiment Analysis of Sentence Embedding Using Machine Learning Techniques. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; pp. 493–496. [CrossRef]