

# Conceptual Map Assessment Through Structure Classification

Laís P. V. Vossen<sup>1,\*</sup>, Isabela Gasparini<sup>1</sup>, Elaine H. T. Oliveira<sup>2</sup>, Lukas Menzel<sup>3</sup>, Sebastian Gombert<sup>4</sup>, Knut Neumann<sup>5</sup> and Hendrik Drachsler<sup>3,4</sup>

<sup>1</sup>*Santa Catarina State University, Joinville, Brazil*

<sup>2</sup>*Federal University of Amazonas, Manaus, Brazil*

<sup>3</sup>*Studiumdigitale & Faculty of Computer Science, Goethe University, Frankfurt am Main, Germany*

<sup>4</sup>*DIPF| Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany*

<sup>5</sup>*IPN| Leibniz Institute for Science and Mathematics Education, Kiel, Germany*

## Abstract

Due to their versatility, conceptual maps are used in various educational settings and serve as tools that enable educators to comprehend students' knowledge construction. An essential component for analyzing a concept map is its structure, which can be categorized into three distinct types: spoke, network, and chain. Understanding the predominant structure in a map offers insights into the student's depth of comprehension of the subject. Therefore, this study examined 511 distinct conceptual map structures, classifying them into one of the three types, and used statistical and descriptive information from the maps to train multiclass classification models. As a result, we achieved an 88% accuracy in classification using a Random Forest. This promising outcome can be employed in conceptual map assessment systems to provide real-time feedback to the student.

## Keywords

Concept maps, Random forest, Structure classification

## 1. Introduction

Concept Maps (CM) are a robust graphical tool for organizing and representing knowledge, and they can be used for conducting assessments [1], [2], or even as a study aid that facilitates the observation of knowledge and potential concepts that the student may be struggling to assimilate [3], either way, it is considered a robust test of knowledge and understanding of a topic [4]. The assessment methods for CM are diverse, considering both qualitative and quantitative aspects of

---

*EvaLLAC'24: Workshop on Automatic Evaluation of Learning and Assessment Content, July 08, 2024, Recife, Brazil*

\*Corresponding author.

✉ [lais.vossen@gmail.com](mailto:lais.vossen@gmail.com) (L. P. V. Vossen); [isagasp@gmail.com](mailto:isagasp@gmail.com) (I. Gasparini); [elaine@icomp.ufam.edu.br](mailto:elaine@icomp.ufam.edu.br) (E. H. T. Oliveira); [menzel@studiumdigitale.uni-frankfurt.de](mailto:menzel@studiumdigitale.uni-frankfurt.de) (L. Menzel); [sebastian@outlook.de](mailto:sebastian@outlook.de) (S. Gombert); [neumann@leibniz-ipn.de](mailto:neumann@leibniz-ipn.de) (K. Neumann); [drachsler@studiumdigitale.uni-frankfurt.de](mailto:drachsler@studiumdigitale.uni-frankfurt.de) (H. Drachsler)

🌐 <https://www.udesc.br/professor/isabela.gasparini> (I. Gasparini); <https://edutec.science/team/sebastian-gombert/> (S. Gombert);

[https://www.studiumdigitale.uni-frankfurt.de/91614784/Prof\\_Dr\\_Hendrik\\_Drachsler\\_Wissenschaftl\\_Leitung](https://www.studiumdigitale.uni-frankfurt.de/91614784/Prof_Dr_Hendrik_Drachsler_Wissenschaftl_Leitung) (H. Drachsler)

🆔 0000-0002-1725-0618 (L. P. V. Vossen); 0000-0002-8094-9261 (I. Gasparini); 0000-0003-2884-9359 (E. H. T. Oliveira); 0000-0002-9870-1109 (L. Menzel); 0000-0001-5598-9547 (S. Gombert); 0000-0002-4391-7308 (K. Neumann); 0000-0001-8407-5314 (H. Drachsler)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

their construction [5]. In either form, the analysis of these maps can provide valuable insights into the student's level of learning, highlighting strengths and knowledge gaps. One of the assessment techniques for concept maps, proposed by Kinchin et al. [6], involves categorizing them based on their structure and dividing them into three distinct categories: network, chain, or spoke. These structures proved to be a robust way [7, 8] for the educator to visualize the different levels of understanding of a given topic by analyzing the concept map structure that the student created [9, 4].

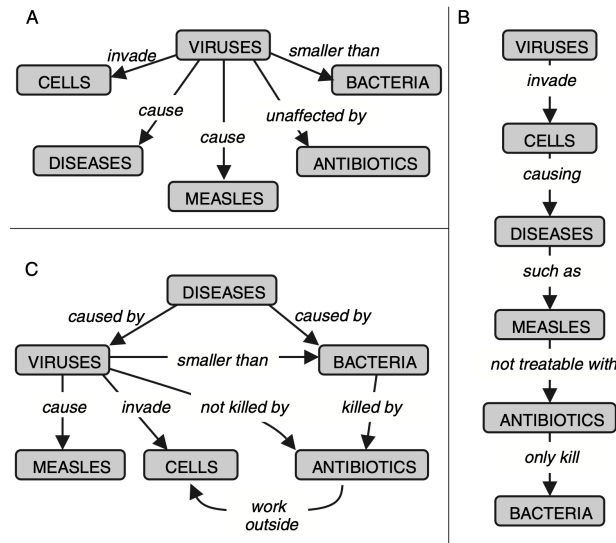
Hence, the relevance of conceptual maps in education is noteworthy, as they serve as a valuable learning instrument that aids educators in the teaching process. Therefore, categorizing the CM by their structure format is crucial to understanding students' rationale, identifying improvement opportunities, and optimizing the learning process. However, the manual categorization process can be time-consuming or even impractical in scenarios with large volumes of produced maps. Therefore, this article addresses the following research question: Can concept maps be classified as Spoke, Chain, or Network based on their features and without relying on human visual interpretation?

To answer this question, we manually classified over 500 hundred CM about two distinct topics in one of the three structures. Through feature engineering, we extracted features that characterize each map and used this information to train various machine learning models to find one capable of identifying the maps into either a network, spoke, or chain structure with satisfactory accuracy.

## **2. Conceptual Maps and their Structural types**

The need to establish standards for evaluating these conceptual maps has arisen. Different evaluation approaches emerged, including quantitative, holistic, similarity-based on a model conceptual map, or qualitative ones [3]. Among these evaluation methods, Kinchin et al. [6] proposed categorizing conceptual map structures into three types: spoke, network, and chain, as illustrated in Figure 1.

An example of a spoke-type map is depicted in Figure 1 in panel A. This type of map exhibits a circular structure in which all the related aspects of the topic are directly linked to the core concept but not directly connected. Such a structure often indicates superficial and underdeveloped knowledge of a subject. Although new knowledge can be more easily integrated, it may only be accessible through reference to the core concept [4, 6]. In the same figure, panel B illustrates a chain-type map. These structures lack linkage between nodes in different hierarchies and tend to move through available information from start to finish quickly. Consequently, understanding concepts in the middle of the sequence without starting from the beginning may be challenging for students. Additionally, incorporating new knowledge, especially if it disrupts an established sequence, can be difficult [4, 6]. Finally, panel C exemplifies a network-type map, distinguished by its robust and adaptable cognitive typology. This structure enables the comprehension of a diverse range of concepts through different but related routes of thought, demonstrating students' flexibility in understanding the concept beyond a connection to a core concept or following a specific line of thought [4, 6].



**Figure 1:** Examples of concept maps in three structures: A) spoke; B) chain; C) network

Source: Kinchin et al. (2005) [7]

### 3. The Database

A collaboration with Goethe University provided the data from the use of Hyperchalk [10], a digital and collaborative whiteboard, for a six-week period in secondary and high schools in the Schleswig-Holstein state, in Germany. Preparation for analysis involved a dataset comprising a total of 511 distinct maps with at least two links between concepts and, therefore, at least three concepts. The maps are divided into two categories, namely 226 concept maps about the school system and 285 concept maps about evolutionism.

#### 3.1. Data preparation and cleaning

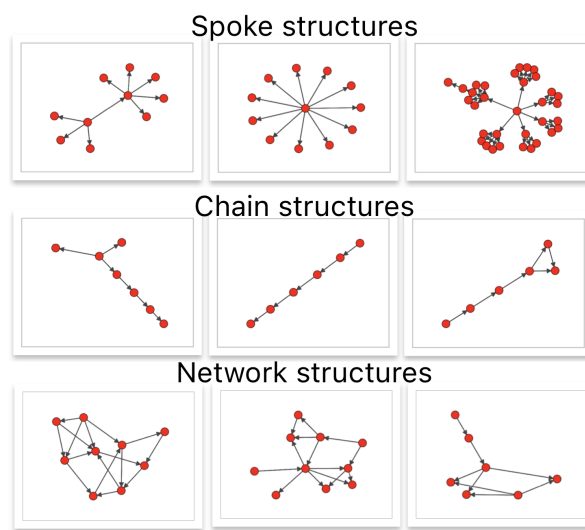
The dataset underwent a transformation process to turn the nodes and edges connections into a directed graph structure, allowing the extraction of statistics data about the number of connections which node had. There was also a feature engineering procedure whereby each map was compiled within a single row of the dataset containing the following information:

- Map identification: the file name that contained the entire concept map;
- Number of nodes: the sum of all nodes in the concept map;
- Number of edges: the sum of all edges in the concept map;
- Ratio between nodes and edges: calculated by dividing the number of edges by the number of nodes in that map. This variable indicates the overall connectivity of a concept map;
- Mean number of edges per node: the mean number of edges that the nodes have, it represents the mean number of connections by node;
- Standard deviation of edges per node: the expected variation from the mean number of connections by node;

- First, second, and third quartile: the number of connections a node has, below which 25%, 50% and 75% of the smallest number of connections by node lie, respectively;
- Maximum number of edges per node: the maximum number of edges connected to a node that happens in the concept map.

### 3.2. Classification among structures types

After transforming the data into graph structures, each map was plotted as a directed graph using the igraph library [11], omitting all labels and thus retaining only the general format of the map, for that reason, the maps' topic didn't influence the classification. Each image was analyzed and classified as a spoke, chain, or network. To accomplish this, we analyzed the most predominant structure of the graph for classification, and at the end of the classification process, 87 CM were classified as chain, 307 as network, and 117 as spoke.



**Figure 2:** Examples of spoke, chain, and network structures found in the maps

Figure 2 shows some examples of those structures found, in which the first row presents three examples of cases considered as spoke-type structures, with the middle one being the most traditional. The images on the right and left exhibit the main characteristic of spoke structures, featuring key concepts to which various other concepts are connected. However, unlike the traditional image, they have more than one key concept. Similar to the traditional structure, the concepts linked to a key concept lack interconnections among themselves. The second row presents some examples of structures considered chains, with the middle one being the closest to the classic chain structure. The figures on the right and left exhibit the same pattern of having a node connected to the others through a single chain connection. Although these structures deviate from the classic one by having nodes with more ramifications, these nodes are few, and the main characteristic of the lack of connection between concepts at different hierarchical levels is maintained. Finally, the third row presents examples of network structures, where highly connected nodes prevail in all cases, without a well-defined hierarchical structure.

In the dataset, there are several examples of networks with varying degrees of connectivity, and structures that do not exhibit a prevalence of chain or spoke characteristics were considered as networks.

## 4. Automatic Concept Map Structure Classification

To address the research question, the prepared dataset was divided into two parts and balanced to have 100 random samples of each class, with 20% of the data used as a validation dataset and the remainder used for model training. Subsequently, five multiclass classification models were selected, as presented in Table 1. The models were implemented using the Scikit-learn library [12], with the default settings for each model. Then, the accuracy was evaluated for each model using cross-validation with the default 5-fold cross-validation method and the F1-Score for each class. Table 1 shows that the Random Forest model outperforms in all the metrics, making it the best choice for this particular case. The Permutation Importance method showed that the five most important features were, in descending order of importance: standard deviation of edges per node, mean number of edges per node, third quartile, maximum number of edges per node, and number of edges.

**Table 1**

Model results for classification in chain, spoke or network structure

Classifier Model	Accuracy	F1-Score Chain	F1-Score Network	F1-Score Spoke
Decision Tree	0.78	0.79	0.75	0.85
Random Forest	<b>0.88</b>	<b>0.91</b>	<b>0.82</b>	<b>0.95</b>
Logistic Regression	0.85	0.78	0.73	0.83
K Neighbors	0.78	0.78	0.65	0.88
Gradient Boosting	0.85	0.88	0.65	0.87

## 5. Discussion

This work addressed the research question presented in Section 1: *Can concept maps be classified as Spoke, Chain, or Network based on their features and without relying on human visual interpretation?* To achieve this, a thorough cleaning and feature engineering process was employed on the available dataset. As a result, various multiclass classification models were tested, with Random Forest achieving the highest accuracy of 88% for classifying a map into the spoke, chain, or network structures.

The results obtained are promising for integrating automated conceptual map evaluation systems, which could improve educators' routines in utilizing these resources. They could also provide real-time feedback to students regarding the structure of their maps, guiding learning toward further depth. A major challenge of this research lies in the manual classification of maps, as we found no robust guideline to differentiate maps that exhibit characteristics of more than one structure. Therefore, in future work, we suggest a thorough discussion of these classification patterns through detailed analysis of the generated structures, as this may further improve the accuracy and F1-Score results obtained by the models.

## Acknowledgments

We thank the project ALICE (Analyzing Learning for Individualized Competence development in mathematics and science Education) initiative. This research is partially supported by CNPq grant 302959/2023-8 (DT2) and 303443/2023-5 (DT2), and FAPESC Edital n° 48/2022 TO n°2023TR000245.

## References

- [1] M. A. Ruiz-Primo, Examining Concept Maps As An Assessment Tool, in: *Concept Maps: Theory, Methodology, Technology Proc. of the First Int. Conference on Concept Mapping* AJ Cañas, JD Novak, FM González, Eds. Pamplona, Spain, 2004.
- [2] M. Strautmane, Concept map-based knowledge assessment tasks and their scoring criteria: An overview, in: *Concept maps: Theory, methodology, technology. Proceedings of the fifth international conference on concept mapping, volume 2, 2012*, pp. 80–88.
- [3] Kirsten E. De Ries, H. Schaap, A.-M. M. J. A. P. Van Loon, M. M. H. Kral, P. C. Meijer, A literature review of open-ended concept maps as a research instrument to study knowledge and learning, *Quality & Quantity* 56 (2022) 73–107. URL: <https://link.springer.com/10.1007/s11135-021-01113-x>. doi:10.1007/s11135-021-01113-x.
- [4] D. B. Hay, I. M. Kinchin, Using concept maps to reveal conceptual typologies, *Education + Training* 48 (2006) 127–142. URL: <https://www.emerald.com/insight/content/doi/10.1108/00400910610651764/full/html>. doi:10.1108/00400910610651764.
- [5] J. Keppens, D. Hay, Concept map assessment for teaching computer programming, *Computer Science Education* 18 (2008) 31–42. ISBN: 0899-3408 Publisher: Taylor & Francis.
- [6] I. M. Kinchin, D. B. Hay, A. Adams, How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development, *Educational Research* 42 (2000) 43–57. URL: <https://www.tandfonline.com/doi/full/10.1080/001318800363908>. doi:10.1080/001318800363908.
- [7] I. Kinchin, D. Hay, Using concept maps to optimize the composition of collaborative student groups: a pilot study, *Journal of Advanced Nursing* 51 (2005) 182–187. URL: <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2648.2005.03478.x>. doi:10.1111/j.1365-2648.2005.03478.x.
- [8] I. M. Kinchin, F. A. A. M. De-Leij, D. B. Hay, The evolution of a collaborative concept mapping activity for undergraduate microbiology students, *Journal of Further and Higher Education* 29 (2005) 1–14. URL: <http://www.tandfonline.com/doi/abs/10.1080/03098770500037655>. doi:10.1080/03098770500037655.
- [9] S. Gerstner, F. X. Bogner, Concept map structure, gender and teaching methods: an investigation of students' science learning, *Educational Research* 51 (2009) 425–438. URL: <https://www.tandfonline.com/doi/full/10.1080/00131880903354758>. doi:10.1080/00131880903354758.
- [10] L. Menzel, S. Gombert, D. Di Mitri, H. Drachsler, Superpowers in the classroom: Hyperchalk is an online whiteboard for learning analytics data collection, in: I. Hilliger, P. J. Muñoz-Merino, T. De Laet, A. Ortega-Arranz, T. Farrell (Eds.), *Educating for a New Future: Making*

Sense of Technology-Enhanced Learning Adoption, volume 13450, Springer International Publishing, 2022, pp. 463–469. URL: [https://link.springer.com/10.1007/978-3-031-16290-9\\_37](https://link.springer.com/10.1007/978-3-031-16290-9_37). doi:10.1007/978-3-031-16290-9\_37.

- [11] G. Csárdi, T. Nepusz, The igraph software package for complex network research, 2006. URL: <https://api.semanticscholar.org/CorpusID:16923281>.
- [12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in python, *J. Mach. Learn. Res.* 12 (2011) 2825–2830.