Identifying Unusual Topic Alignment Patterns in E-Learning Materials

Beata Gancevska, Simona Ramanauskaitė

Vilnius Gediminas Technical University, Faculty of Fundamental Sciences, Department of Information Technologies

Introduction

E-learning courses contain various types of learning materials organized according to course topics. However, the content of these materials does not always correspond clearly to a single topic. Some materials may cover several topics, while others may contain general information or content that is only indirectly related to the course topics. This makes it difficult to evaluate how well learning materials correspond to the structure of a course.

One possible approach to this problem is semantic analysis, which allows the relationship between learning materials and course topics to be examined based on their meaning rather than simply matching keywords. Large language models (LLMs) provide new possibilities for performing such evaluations, especially when dealing with large amounts of textual content.

This research focuses on identifying **unusual semantic alignment patterns between learning materials and course topics**. Each slide is evaluated against all topics within its course, rather than focusing on a single topic. Particular attention is given to **materials with weak alignment, materials strongly related to multiple topics, and differences in topic coverage**. The aim is to explore whether such patterns can help identify learning materials that may require further examination and support the development of recommendations for improving course content organization.

Materials & Methods

Data from **3 Moodle courses** included **25 presentations, 649 slides and 47 course topics**. The full text of each slide was extracted without modification.

ChatGPT (GPT-6 Astra Max and GPT-6.1 Sol Max) was used to assess each slide against every course topic on a **0–100 semantic alignment scale. Two prompting frameworks, TAG and R.A.C.E., were initially tested**. A preliminary review showed that both produced very similar scores. **R.A.C.E. was selected** because it allows a clearer definition of the model's role, evaluation context, and expected output. Slides with insufficient textual information were assigned a score of 0.

An **independent human evaluation** was conducted on **21 slides from 3 courses, covering 329 slide–topic pairs**. The human evaluator assigned scores without seeing the LLM results. **GPT-6 Astra Max** showed stronger agreement with human scores **(MAE = 3.99, r = 0.941**) than GPT-6.1 Sol Max (MAE = 5.07, r = 0.906). **Therefore, GPT-6 Astra Max was selected for further analysis.**

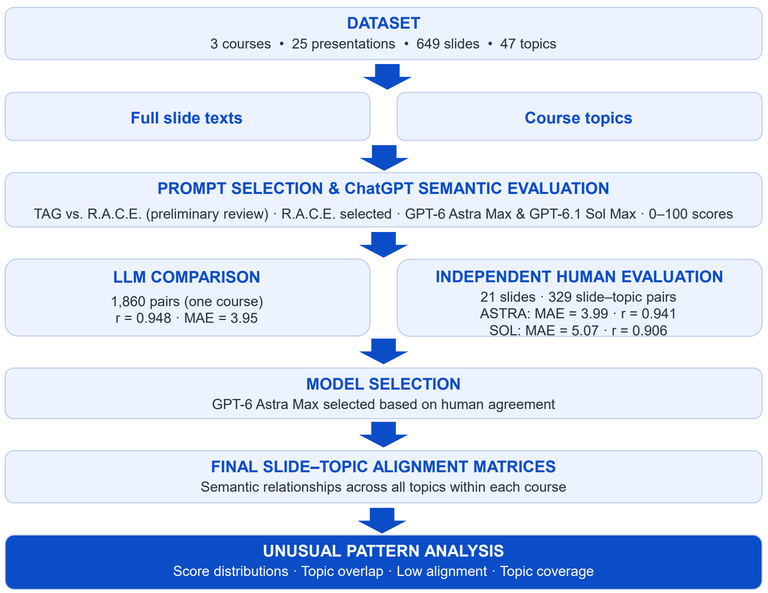

Fig. 1. Overview of the semantic alignment evaluation process

Results

The analysis showed differences in semantic alignment patterns across the three courses. Course 1 had the highest share of strong topic overlap (37.9%), while Course 2 had the highest share of slides with no alignment (27.0%). Multi-topic alignment was common across all three courses, ranging from 46.0% to 62.1% (*Fig. 2*).

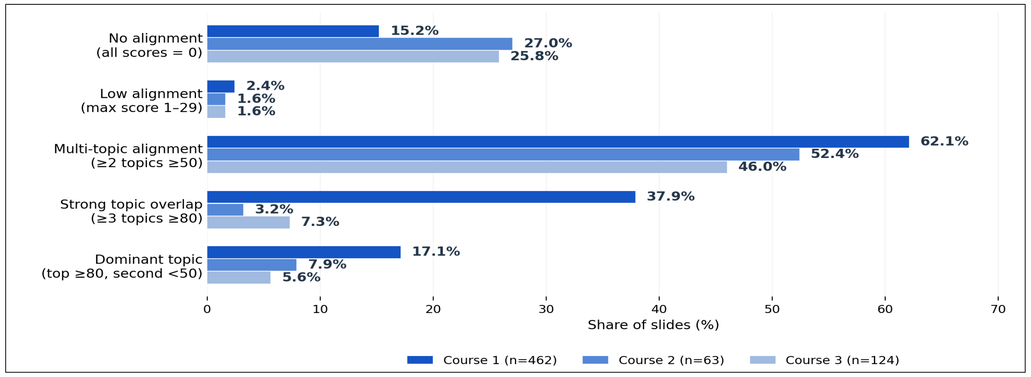

Fig. 2. Five semantic alignment patterns across three courses.

Selected slides show no, weak, dominant and multi-topic alignment (*Fig. 3*).

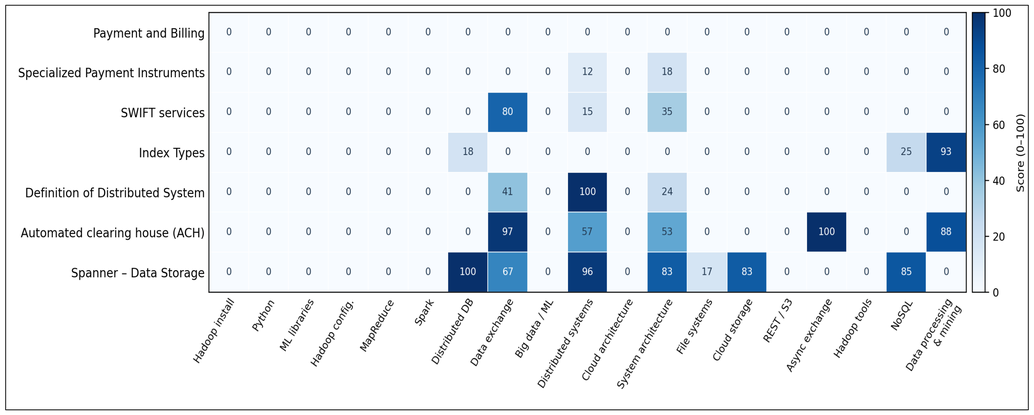

Fig. 3. Slide–topic alignment heatmap for seven selected Course 1 slides.

Three topics had no strongly aligned slides (≥80), while Distributed Systems had 178 (*Fig. 4*).

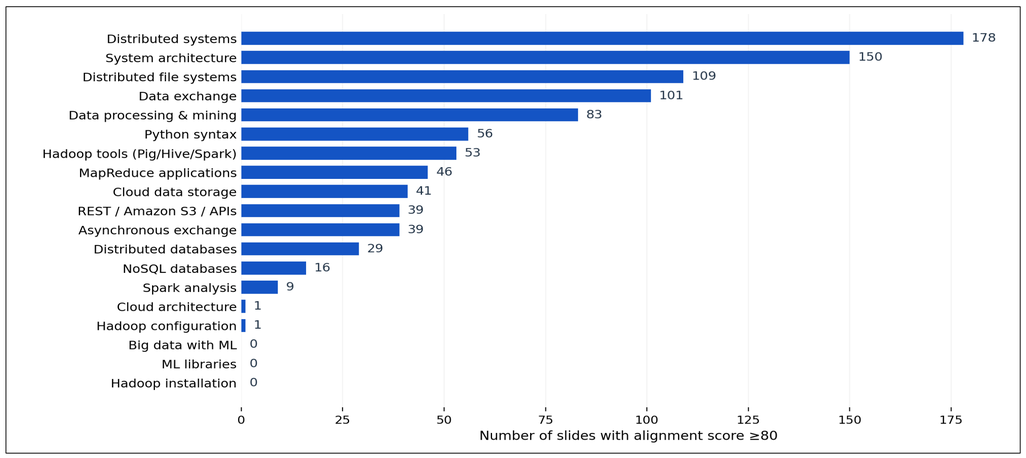

Fig. 4. Course 1 topics ranked by the number of strongly aligned slides (score ≥80).

Interpretation of Results

The results showed differences in semantic alignment across courses and topics. **Low alignment scores** may be caused by limited slide text or content that is not closely related to the course topics. **High scores for multiple topics** may indicate that the learning material covers several related concepts rather than being incorrectly organized.

Some course topics had **many strongly aligned slides, while others had very few or none**. These differences may help identify areas that need further review and **support the development of recommendations for improving course content organization**. However, **alignment scores alone are not enough to determine whether learning materials are properly organized**.

Conclusion and Future Work

This study explored the use of LLMs to evaluate the semantic alignment between e-learning materials and course topics. The results showed different alignment patterns, including **weak relationships with course topics, strong alignment with multiple topics, and differences in topic coverage**. These patterns may help identify learning materials that require further examination and **support the development of recommendations for improving course content organization**.

However, the study is limited to **three courses**, and the validation of LLM-generated scores is based on a **small sample evaluated by one human evaluator**. Additionally, the analysis focuses on the textual content of presentation slides, **without considering visual elements** that may provide important contextual information.

Future research should **expand the analysis to more courses, involve additional human evaluators, and further investigate the identified unusual patterns**. This could help improve the reliability of semantic alignment evaluation and support the development of **automated recommendations for organizing e-learning materials**.

Core References

1. Doostmohammadi, E., Holmström, O., & Kuhlmann, M. (2024). How reliable are automatic evaluation methods for instruction-tuned LLMs? In *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 6321–6336). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-emnlp.367
2. He-Yueya, J., Goodman, N. D., & Brunskill, E. (2024). Evaluating and optimizing educational content with large language model judgments. In *Proceedings of the 17th International Conference on Educational Data Mining*. https://doi.org/10.5281/zenodo.12729776
3. Kwak, Y., & Pardos, Z. A. (2024). Bridging large language model disparities: Skill tagging of multilingual educational content. *British Journal of Educational Technology, 55*(5), 2039–2057. https://doi.org/10.1111/bjet.13465
4. Molavi Hajiagha, M., Moein, M., Tavakoli, M., Faraji, A., Kismihók, G., & Mol, S. T. (2026). Embedding-based rankings of educational resources based on learning outcome alignment: Benchmarking, expert validation, and learner performance. In *Proceedings of the LAK26: 16th International Learning Analytics and Knowledge Conference* (pp. 829–836). Association for Computing Machinery. https://doi.org/10.1145/3785022.3785073
5. Pereira, E., Nsair, S., Pereira, L. R., & Grant, K. (2024). Constructive alignment in a graduate-level project management course: An innovative framework using large language models. *International Journal of Educational Technology in Higher Education, 21*, Article 25. https://doi.org/10.1186/s41239-024-00457-2