Research on Different AI Approaches to Measuring Curricular Overlap Across University Programs*

Simona Ramanauskaitė[1,*,†], Pavel Stefanovič[1,†], Eglė Baradinskienė[2,†], Cristina Vărvăruc[3,†], Muhammad Imran[4,†], Darja Bubnova[5,†], Cristian Maximiliano Rodriguez Rivero[6,†], Tõnis Raamets[7,†] and Renata Karbauskienė[1,†]

[1] *Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223, Vilnius, Lithuania*
[2] *Exacaster, J. Basanavičiaus g. 26, LT-03224, Vilnius Lithuania*
[3] *Alexandru Ioan Cuza University of Iasi, Bulevardul Carol I 11, 700506, Iași, Romania*
[4] *Fondazione Bruno Kessler, Via Sommarive, 18, 38123 Povo TN, Italy*
[5] *Riga Technical University, Āzenes iela 12, Kurzemes rajons, LV-1048, Rīga, Latvia*
[6] *Universitat Politècnica de Catalunya, Av. de les Bases de Manresa, 61, 73, 08242 Manresa, Barcelona, Spain*
[7] *Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia*

Abstract

University curricula's increasing complexity and diversity necessitate efficient methods for assessing course similarities across institutions. This research investigates and compares traditional text-similarity techniques and generative AI (GenAI) approaches for estimating curricular overlap among master's programs in computer science at four European universities. A manually curated dataset of course similarities serves as the ground truth for evaluating automated methods. Traditional approaches employed SBERT-based embeddings and cosine similarity to analyze course topics, achieving a mean absolute error (MAE) of 4%. In contrast, GenAI models (GPT-4 and GPT-5) demonstrated overoptimistic similarity estimations, with an MAE of 31% and a lower correlation manual assessments than the traditional approach. While GenAI offers valuable qualitative insights and a justification for similarities, its quantitative reliability remains limited. The findings highlight that GenAI tools are still lagging behind specifically crafted NLP solutions in this task, but could be used as supplemented tool.

Keywords

course, similarity, embedding, cosine, generative AI, university.

1. Introduction

In recent years, the rapid expansion and diversification of university programs have made it increasingly important to understand how curricula align, diverge, or overlap across institutions Curricular overlap (the extent to which courses, topics, and learning outcomes are shared between programs) plays a critical role in academic benchmarking, credit transfer, accreditation, and strategic program design. However, traditional methods for assessing curricular similarity, such as manual syllabus comparison or keyword analysis, are time-consuming, subjective, and often unable to capture nuanced relationships between disciplines.

Advances in artificial intelligence (AI) and natural language processing (NLP) now offer promising new avenues for automating and improving curriculum analysis. From semantic embeddings that model textual similarity to clustering algorithms that group related learning outcomes, AI-driven approaches can provide deeper insights into academic programs' structure and

*IVUS 2025: Information Society and University Studies 2024, May 15, Kaunas, Lithuania
* Corresponding author.
† These authors contributed equally.
simona.ramanauskaite@vilniustech.lt (S. Ramanauskaitė); pavel.stefanovic@vilniustech.lt (P. Stefanovič); egle@exacaster.com (E. Baradinskienė); cristina.bulat@digital-innovation.zone (C. Vărvăruc); mimran@fbk.eu (M. Imran); Darja.Bubnova@rbs.lv (D. Bubnova); cristian.maximiliano.rodriguez@upc.edu (C. M. Rodriguez Rivero); tonis.raamets@taltech.ee (T. Raamets); renata.karbauskiene@vilniustech.lt (R. Karbauskienė)
0000-0003-3195-4280 (S. Ramanauskaitė); 0000-0003-0503-4157 (P. Stefanovič); 0000-0003-4367-0197 (C. Vărvăruc); 0000-0002-4315-7476 (C. M. Rodriguez Rivero)

content. Yet, despite these opportunities, there remains limited comparative research evaluating how different AI techniques perform at measuring curricular overlap. This is particularly true for different study institutions.

This study aims to address that gap by systematically investigating and comparing various AI methodologies for quantifying curricular overlap among university programs. Specifically, it examines machine learning, deep learning, and hybrid NLP approaches to determine their relative accuracy, scalability, and interpretability in curricular mapping tasks.
At the same time, these research questions are raised in this research:

1. How capable are generative AI (GenAI) tools to estimate the course overlap level, based on the course syllabus?
2. How accurate are traditional text similarity estimation methods and application solutions for course overall estimation?

The findings seek to contribute to both educational research and institutional practice by offering evidence-based recommendations for leveraging AI to enhance curricular transparency, alignment, and innovation in higher education.

2. Related Work

Research into understanding how courses or programs overlap, align or diverge within or across institutions has gained traction in recent years. This is particularly true as institutions examine credit transfer, articulation, on curricular transparency, and program design. The initial motivation was more oriented to student performance prediction [1]. Application of learning objectives via TF-IDF, Jaccard, and semantic similarity in a grade-prediction context demonstrated that course similarity estimation is feasible and relevant to downstream analytics. Automated course outcome mapping, using a modified “Barclay’s coefficient”, was proved to be effective in personalized learning path generation [2].

The need to concentrate on the curriculum, course overall is relevant in different studies. Yan Zhou et al. [3] highlights the challenge of acquiring similar learning outcomes across four parallel thematic learning communities in a medical undergraduate curriculum. This helps manage curricular overlaps and divergences in learning outcomes across four program variants. In this way course similarity estimation can be used not only for curriculum design, but for teaching and content recommendations as well [4]. The recommendations usually are based on integrating both course descriptions and well as other data, reflecting study quality and places for improvement.

The problem of course content matching is relevant as well. H. Shahd, I. Salma and H. Amjad [5] use an ontological comparison-based approach, where each course is transformed into an ontology and then the course similarity is estimated based on the ontological representation and deep learning similarity scoring. Meanwhile X. Yinuo and Z. A. Pardos extracted course similarity signals using subword embeddings, finding out FastText and SBERT integration works best for equivalency validation. Student thesis “Text analysis techniques and determining the similarity of course contents” [7] summarizes different approaches that can be used for course similarity estimation. Course similarity estimation with artificial intelligence solutions is just one of the cases how higher education can utilize AI [8]. Meanwhile generative AI application in higher education is still controversial [9, 10, 11].

These studies illustrate that work has been done on course similarity mapping, learning outcomes, curriculum recommendations, and program structure analytics. However, several gaps remain.

1. Many studies focus on single institutions or single programs rather than cross-institutional or even cross-boarded study programs.
2. A range of similarity metrics (TF-IDF, Jaccard, embeddings) are used but often without systematic comparison of different AI approaches in the context of curricular overlap.

3. Generative AI is starting to be used for course content development, students' open-ended answer evaluation, however it is still not popular for study program administration, and course overlap estimation.

Thus, this research, which compares different AI methods for quantifying curricular overlap across programs, is timely and fills a meaningful gap.

3. Research Approach for Course Similarity Estimation

3.1. General Approach of the Research

The research was designed to check the similarity between courses in several university study programs. All courses in the selected 4 master's study programs were presented as PDF files. This was done by presenting each of them as a separate file and combining them into folders, according to the institution. To get the ground truth of course similarity scores, manual course similarity estimation was done (see Figure 1). At the same time courses at each institution were divided into testing and validation datasets. The testing dataset was shared (dotted line in Figure 1) with the research team to be used in a traditional text-similarity approach and in parallel with GenAI application research. The dataset was used to adjust the prompts' or threshold values, achieving better similarity accuracy. Meanwhile the validation dataset (dot-dash lines in Figure 1) was applied to get the final results for comparison.

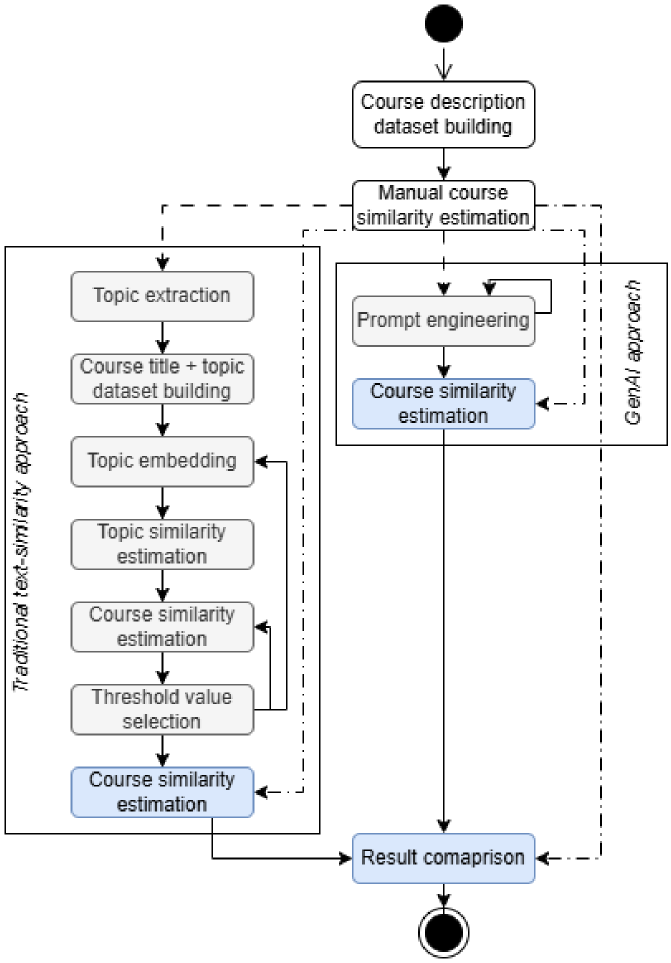


Figure 1: The main flow of the experiment (generated by authors).

The main parts of this research schema are defined further, concentrating on what actions, methodologies and approaches were used in each of them.

3.2. Dataset Building for Course Similarity Estimation

For the research master’s study programs courses were gathered from 4 universities: Riga Technical University, RTU (study program “Management of Smart, Resilient, and Interconnected Systems”); Universitat Politècnica de Catalunya, UPC (study program “Master of Machine Learning and Cybersecurity for Internet Connected Systems”); Tallinn University of Technology, TalTech (study program “Industrial Engineering and Management”); Vilnius Gediminas Technical University, VILNIUS TECH (study programs “Engineering of Artificial Intelligence” and “Management of Artificial Intelligence Solutions”). Each university individually selected courses as candidates for Collaborative Online International Learning (COIL) among the universities. One third was selected for validation and two thirds were chosen for testing purposes (see Figure 2).

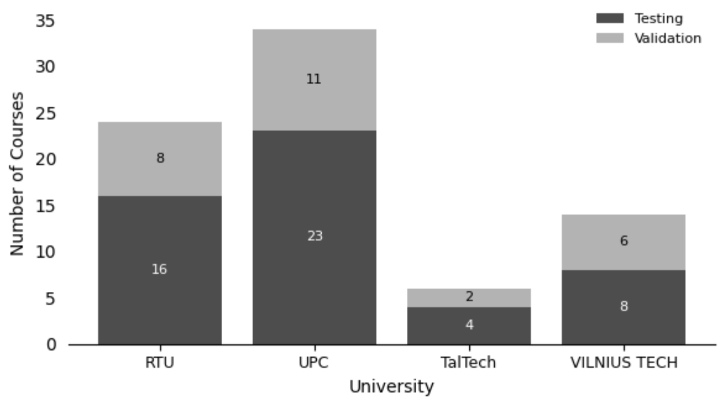


Figure 2: Distribution of courses of each university for testing and validation (generated by authors).

The course selection for testing or validation was done by experts. This was to guarantee that both similar and not similar courses at different universities would fall in both testing and validation sets. Therefore, the division of the set was done after manual course similarity estimation.

Each university estimated the course similarity by a course owner (the main teaching professor) who was responsible for revising all the other courses of the other 3 universities and marking the match of each of the courses to its course, expressing it in percentage. In the next step, an analysis of matching was done, to estimate a two-directional relationship between the mapped courses. For example, if course CA in institution IA was mapped to course CB in institution IB, did institution IB also map course CB to course AC? The possible mismatches were revised and discussed in a group session between all 4 universities. This approach was used to guarantee mapping accuracy.

Despite the fact that all courses were from the computer science field, the specifics of each study program affected the relatively low course coverage as all course topics had to be taken into account. Very aggregated statistics for each university's course coverage are the following:

1. RTU had 3 courses, which were not covered by any of the partner university courses, while the maximum match with existing courses was 50%, while usually the similarity was 10-15%.
2. UPC had 12 courses, which were not covered by any of the partner university courses, while the maximum match with existing courses was 90%, while usually the similarity was 10%.
3. TalTech had 2 courses, which were not covered by any of the partner university courses, while the maximum match with the existing course was 40%, while usually the similarity was 15%.
4. VILNIUS TECH had all courses covered by some of the partner university courses, while the maximum match with existing courses was 50%, while usually the similarity was 10-20%.

To implement this mapping between all institutions, a big number of course comparisons had to be done. Including each course owner, the task was distributed, but it took about 2 months to finish it. This just highlights the need for an automated approach, to reduce workload.

3.3. Methodology Behind Traditional Text-Similarity Approach

To apply traditional text-similarity estimation methods, the obtained courses were slipped into lists of topics, which must be covered in the course (see Figure 1). In this way, further analysis used a dataset, where each row was composed of the course title and the topic title. This eliminated all the remaining data and concentrated on English language only (some universities provide course descriptions both in English and national languages).

All the text was embedded using 4 different SBERT models:

1. all-MiniML-L6v2 – 384 dimensions.
2. all-mpnet-base-v2 – 768 dimensions.
3. allenai-specter – 768 dimensions.
4. paraphrase-distilroberta-base-v1 – 768 dimensions.

Traditional Bag-of-words and ngram methods were tested and rejected as potential to be used because of different terminology in the course topics and lack of semantic similarity estimation.

With the embeddings a similarity estimation was executed between each pair of universities, including all courses and topics. The experiments included these similarity estimation methods:

1. Cosine similarity.
2. Euclidean distance.
3. Manhattan distance.
4. Chebyshev distance.

The similarity between each topic was logged, but the matrix was filtered, leaving only the matches, which are higher than the threshold value. This decision was made to find topics between two courses that match on a sufficiently high level. This would mean these topics could be used for COIL. Meanwhile for course similarity, the selected approach takes the maximum similarity for the topic, and if it reaches the threshold value, it is taken as the similarity of the topic, while if the similarity score is too low, the similarity was set to 0 and the final course similarity was calculated as the average of all the topics' similarities.

Experiments using different embedding models, similarity estimation methods and threshold values were executed to find a combination, where the Mean Absolute Error (MAE) was the smallest. This metric was selected taking into account course similarity, which always ranges between 0 and 100%.

3.4. Methodology Behind GenAI Approach

For the GenAI approach the full course description was used. The comparison was done among the validation courses, while among the testing courses, an initial set of experiments was executed to find the most appropriate prompt for this task. For it, two files from the compared courses were provided and then according to the CIDI framework [12], a prompt was generated.

The final prompt used for the experiment was adjusted to be the following, considering that the topics of the courses might vary from AI to cybersecurity, the internet of things, management and others (see Figure 3).

The prompt was tested with ChatGPT (GPT-4 and GPT-5) to adjust the prompt. All the prompt engineering was done by an external institution, to ensure none of the universities would affect the prompt to adjust to its specifics. At the same time the prompt was formed to provide additional data that could be used to support the similarity score. The score was useful both in the prompt engineering phase as well as giving additional trust in the result during its application in the validation course.

Context:
You are a professor at a university responsible for evaluating course equivalencies. Your task is to compare courses across different universities to determine whether they are similar or not, based on key academic criteria and what proportion of the second course can be reused in the first course.
Criteria:
1. Course content: Specific topics the student studies or works on in the course.
2. Condition: Courses can be considered similar if they have: A majority of the topics are aligned. If they diverge significantly in any of these areas, the courses should be considered not similar.
Task:
1. You are provided two course descriptions, each in a separate file.
2. Compare the courses based on the course topics.
3. Write the comparison in short but straightforward aspects.
4. Clearly state if the courses are similar or not AND provide a concise justification based on the criteria and condition.
5. Don't be too focused on details when comparing.
6. Provide the overall similarity % on how the first course is similar to the second one.

Figure 3: Example prompt for comparison of two courses (generated by authors).

4. Course Overlap Estimation Research Results

4.1. Insights from the Manual Course Overlap Analysis

The executed experiment revealed not only differences between traditional text-similarity and GenAI approaches, but also paved a way for better understanding of the analyzed study programs. One of the key findings from the manual course similarity estimation is that the course title cannot be used as the main criteria for course coverage analysis. As an example, two universities had a course with an identical title "Fundamentals of Artificial Intelligence", but the course owners defined the similarity between these courses as 5% and 10%. This was affected by different understanding of what is AI – one university concentrated on knowledge presentation, search algorithms, graph theory, logic, Markov chains. In contrast, another university dedicated attention to data collection and cleaning, supervised and unsupervised learning, AI model evaluation, metrics, data, results visualization.

But at the same time, the course title is relevant. The context of the course might be reflected in the title, while course topics could be difficult to explain without a title. For example, in a course "Machine learning for cybersecurity" topics like "Patterns in high-dimensional data", "Unsupervised discovery of rare events" represent the machine learning topic, but do not present anything specific to cybersecurity. In a course "Legal Aspects of Digitalization", topics like "The European Union legal system" or "Contracts" present the legal aspects, but do not address digitalization as the background of the topic. This is why in the applied approaches both titles and topics are included.

Another trivial, but confirmed result is that course overlap estimation is a two directional process. It means both universities have to evaluate how their courses compare with others. If course CA has a very high match with course CB, not necessarily course CB will have a high match with course CA. Especially it's relevant to different size courses – one course could cover a whole other course and have additional topics to it. But it's relevant to the same size course too. Some courses concentrate on a wide topic, mentioning many different technologies, whereas other courses are very narrow and dig deeper into the topic, analyzing various aspects of it.

A fragment of the course overlap matrix for two university courses is presented in Figure 4, where rows indicate the course for which we are looking at another course, which could cover certain parts of it and estimate how much the course integration could cover the planned material.

	Artificial intelligence driven process control	Business applications of artificial intelligence	Business operations development and management	IT quality assurance	Legal aspects of digitalization	Business analytics	Artificial intelligence in healthcare	Decision making and probabilistic reasoning
Neural networks							10%	
AI project engineering		10%	10%	5%	10%			
Bussiness analytics			10%			30%		
Fundamentals of AI	10%	10%					30%	10%

Figure 4: Fragment of manual course overlap estimation (generated by authors).

It was noticed that all course owners used the 5% step for similarity estimation and provided no data if the courses had less than 5% overlap.

4.2. Results of the Traditional Text-Similarity Approach

The initial experiments with text tokenization approaches using Bag-of-Words and ngrams indicated the topic similarity estimation is not accurate as the semantics of the key terms is not taken into account, while the supporting transition words gain high importance. Therefore, the experiments switched to work on embedding usage, using different SBERT models.

The smallest SBERT model was all-MiniML-L6v2. Its application to topic similarity estimation generates very small similarity scores. Taking into account the possible error ranges and the small similarity scores, SBERT models with 768 dimensions were used. Despite the all-mpnet-base-v2 was of general purpose, while allenai-specter was oriented on scientific/research texts (what is similar to course descriptions in masters study courses) and paraphrase-distilroberta-base-v1 potentially could be more accurate on paraphrased texts, finally the more comprehensive all-mpnet-base-v2 was selected, because of its ability to achieve higher accuracy for course topic comparison.

Meanwhile, the application of different vector similarity estimation approaches leads to the selection of a cosine similarity metric as best matching our task. It indicates the highest correlation between manually mapped topics and the similarity of their embeddings. This could be explained by the fact that the direction of SBERT embeddings encodes a meaning more than their magnitude as the embedding has a needed level of dimension and the course topics are from a wide range of topics.

The last phase was oriented on how to tune-up the cosine similarity scores for each topic to course similarity, to achieve the smallest Mean Absolute Error (MAE). The highest accuracy was achieved by using a threshold value of 0.5 for the course title and topic cosine similarity, embedded with all-mpnet-base-v2 SBERT model. The course similarity values, for the same fragment as in Figure 4 are presented in Figure 5.

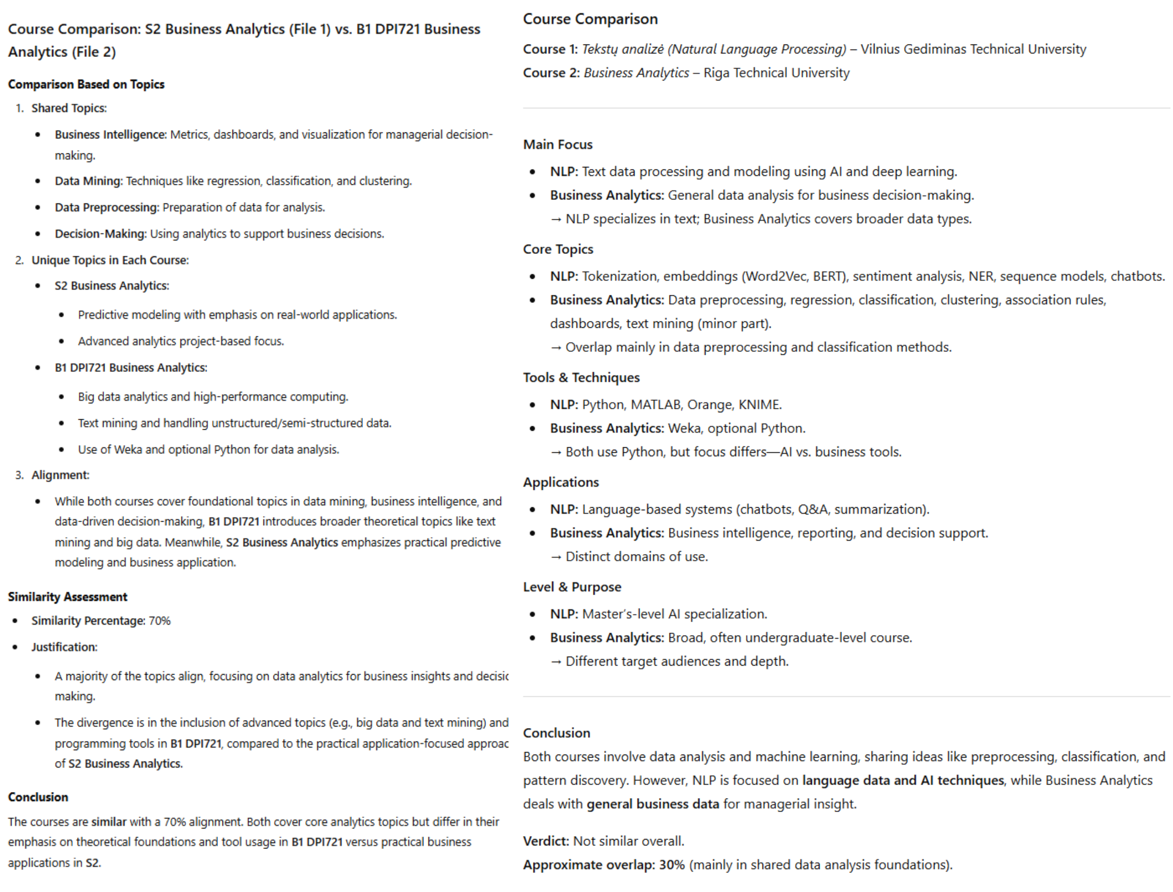
	Artificial intelligence driven process control systems	Business applications of artificial intelligence	Business operations development and management	IT quality assurance	Legal aspects of digitalization	Business analytics	Artificial intelligence in healthcare	Decision making and probabilistic reasoning
Neural networks							6%	
AI project engineering		5%	2%	4%	2%			
Bussiness analytics			4%			50%		
Fundamentals of AI	6%	3%					16%	3%

Figure 5: Fragment of traditional text-similarity approach results for course overlap estimation (generated by authors).

Analyzing the MAE value, the difference between the manually and traditional approach, based on text-similarity, percent is 4% (the standard deviation is 6). This score is quite promising as we could expect the similarity between the course owner and the automatic approach to be within a 0-15% difference. This is not critical, considering some variations can be invisible looking at the topic list. At the same time, by analyzing courses that have at least one topic, reaching the threshold value and accepting the course owners' opinion that there is at least one similarity; the prediction is matched with 92% accuracy. In most cases, it's false-positive case, indicating that the automated approach will provide more candidates for mapping. This could lead to a solution, where the automated approach does the initial mapping and then from the selected candidates manual mapping can be done for higher accuracy.

4.3. Results of the GenAI Approach

As a result of prompt engineering, the defined prompt returns not only a course similarity score, but also a justification for the decision. Figure 6 presents the results of GPT-4 on the left and GPT-5 on the right.


Course Comparison: S2 Business Analytics (File 1) vs. B1 DPI721 Business Analytics (File 2)

Comparison Based on Topics

1. Shared Topics:
   - **Business Intelligence:** Metrics, dashboards, and visualization for managerial decision-making.
   - **Data Mining:** Techniques like regression, classification, and clustering.
   - **Data Preprocessing:** Preparation of data for analysis.
   - **Decision-Making:** Using analytics to support business decisions.
2. Unique Topics in Each Course:
   - S2 Business Analytics:
     - Predictive modeling with emphasis on real-world applications.
     - Advanced analytics project-based focus.
   - B1 DPI721 Business Analytics:
     - Big data analytics and high-performance computing.
     - Text mining and handling unstructured/semi-structured data.
     - Use of Weka and optional Python for data analysis.
3. Alignment:
   - While both courses cover foundational topics in data mining, business intelligence, and data-driven decision-making, **B1 DPI721** introduces broader theoretical topics like text mining and big data. Meanwhile, **S2 Business Analytics** emphasizes practical predictive modeling and business application.

Similarity Assessment

- Similarity Percentage: 70%
- Justification:
  - A majority of the topics align, focusing on data analytics for business insights and decisic making.
  - The divergence is in the inclusion of advanced topics (e.g., big data and text mining) and programming tools in **B1 DPI721**, compared to the practical application-focused approac of **S2 Business Analytics**.

Conclusion

The courses are **similar** with a 70% alignment. Both cover core analytics topics but differ in their emphasis on theoretical foundations and tool usage in **B1 DPI721** versus practical business applications in **S2**.

Course Comparison

Course 1: *Tekstų analizė (Natural Language Processing)* – Vilnius Gediminas Technical University
Course 2: *Business Analytics* – Riga Technical University

Main Focus

- **NLP:** Text data processing and modeling using AI and deep learning.
- **Business Analytics:** General data analysis for business decision-making.
  → NLP specializes in text; Business Analytics covers broader data types.

Core Topics

- **NLP:** Tokenization, embeddings (Word2Vec, BERT), sentiment analysis, NER, sequence models, chatbots.
- **Business Analytics:** Data preprocessing, regression, classification, clustering, association rules, dashboards, text mining (minor part).
  → Overlap mainly in data preprocessing and classification methods.

Tools & Techniques

- **NLP:** Python, MATLAB, Orange, KNIME.
- **Business Analytics:** Weka, optional Python.
  → Both use Python, but focus differs—AI vs. business tools.

Applications

- **NLP:** Language-based systems (chatbots, Q&A, summarization).
- **Business Analytics:** Business intelligence, reporting, and decision support.
  → Distinct domains of use.

Level & Purpose

- **NLP:** Master's-level AI specialization.
- **Business Analytics:** Broad, often undergraduate-level course.
  → Different target audiences and depth.

Conclusion

Both courses involve data analysis and machine learning, sharing ideas like preprocessing, classification, and pattern discovery. However, NLP is focused on **language data and AI techniques**, while Business Analytics deals with **general business data** for managerial insight.

Verdict: Not similar overall.
Approximate overlap: 30% (mainly in shared data analysis foundations).

Figure 6: Example of GenAI output for two course comparisons (generated by GPT-4 on the left and GPT-5 on the right).

The biggest value of this approach is the listing of summarized information on what is common between these two courses and where the main differences are. However, the main issue with this approach (despite the use of the GPT version) is overly optimistic evaluation of course similarity. In the examples provided in Figure 6, the right example illustrates 30% of the course matches, however the "Natural Language Processing" course was not mapped with the "Business Analytics" course by the course owner. Meanwhile, compared two "Business Analytics" courses in different universities, they had similar topics (the course owner marked it as 50% coverage for the first course), but GenAI approaches provided 70% (in GPT-4) or even 73% (in GPT-5) of similarity, which is too optimistic.

The over optimistic evaluation of course similarity is visible in Figure 7, as GenAI always indicated course similarity and it was not lower than 20%. Analyzing the MAE value, the difference between manually and automatically generated course similarity percent is 31% (standard deviation is 9).

	Artificial intelligence driven process control systems	Business applications of artificial intelligence	Business operations development and management	IT quality assurance	Legal aspects of digitalization	Business analytics	Artificial intelligence in healthcare	Decision making and probabilistic reasoning
Neural networks	35%	30%	20%	20%	20%	20%	40%	40%
AI project engineering	30%	40%	55%	25%	35%	30%	30%	35%
Bussiness analytics	30%	35%	60%	25%	20%	70%	20%	40%
Fundamentals of AI	40%	40%	30%	20%	40%	45%	75%	45%

Figure 7: Fragment of GenAI approach (with GPT-4) results for course overlap estimation (generated by authors).

Even after modifying all the proposed values, by reducing them (by 20%), the MAE remained higher than 10%. Reducing the generated similarity score by a different value was not meaningful, as the correlation between the manual and GenAI generated similarities (0.52) was lower than the one obtained by comparing traditional text-similarity approach with manual mapping (0.75). In this way it not only gives too many false-positive results, but the variability is higher.

5. Conclusions

Full course overlap analysis requires a great deal of manual work and expertise in cases where multiple study programs with a big number of courses have to be compared. In this experiment master's study program courses were analyzed, whose duration was up to 2 years. In the case of bachelor's studies, the number of courses and diversity would be even higher, leading to even more manual work. Therefore, if we want to reduce manual workload, an automated or at least partially automated approach has to be used to filter out irrelevant cases.

Implementation of a traditional text-similarity approach, which uses SBERT for topic embedding, cosine similarity for topic similarity estimation and internal logic for course overlap estimation based on topic similarity, allowed us to achieve relatively high course overlap estimation accuracy. In 92% of mappings, the mean average error is less than the expert course overlap step - 5. The remaining 8% of the errors are false positives. This means 84% of the course owner workload can be reduced, as this approach identifies only 16% of compared course pairs, which have at least one topic, reaching the defined threshold value.

Meanwhile the GenAI approach, where GPT-4 and GPT-5 were used, provided at least 20% course overlap between each course combination, leading to overoptimistic course overlap estimation. For the semi-automated approach, where GenAI estimates the initial scores and then course owners revise the matching courses, the workload would not be reduced, but all cases would be marked as at least partly matching. It is nevertheless possible to standardize course data presentation for course owners despite the overly optimistic overlap result.

Acknowledgements

The research results were achieved in European Union co-funded project MERIT under grant agreement No. 101083531.

Declaration on Generative AI

During the preparation of this work, the authors used Wordtune in order to: Improve writing style. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the publication's content.
The authors also used GPT-4 and GPT-5 as systems for the automated course overlap estimation. These tools were part of the research, while was not used for the paper preparation.

References

[1] Li, Jiawei, S. Supraja, Wei Qiu, and Andy WH Khong. "Grade Prediction via Prior Grades and Text Mining on Course Descriptions: Course Outlines and Intended Learning Outcomes." International Educational Data Mining Society (2022).

[2] Huang, J. "Research on the optimization pathways of University Education and teaching reform based on OBE concepts." Applied Mathematics and Nonlinear Sciences 9, no. 1 (2024).

[3] Zhou, Yan, Thomas H. Wieringa, Jasperina Brouwer, Agnes D. Diemers, and Nicolaas A. Bos. "Challenges to acquire similar learning outcomes across four parallel thematic learning communities in a medical undergraduate curriculum." BMC Medical Education 23, no. 1 (2023): 349.

[4] Li, Y., D. Shao, and H. Sun. "Optimizing higher education curriculum design and teaching content recommendation system using data mining algorithms." Applied Mathematics and Nonlinear Sciences 9, no. 1 (2024): 20242427.

[5] Hamdan, Shahd, Salma Ibrahim, and Amjad Hawash. "Measuring the Matches Between University Courses: Ontological Comparison-Based Approach." In 2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), pp. 351-358. IEEE, 2024.

[6] Xu, Yinuo, and Zach A. Pardos. "Extracting course similarity signal using subword embeddings." In Proceedings of the 14th Learning Analytics and Knowledge Conference, pp. 857-863. 2024.

[7] Östensson, Emilia. "Text analysis techniques and determining the similarity of course contents." (2025).

[8] Retscher, Guenther. "Exploring the intersection of artificial intelligence and higher education: opportunities and challenges in the context of geomatics education." Applied Geomatics 17, no. 1 (2025): 49-61.

[9] Ullmann, Thomas, Chris Edwards, Duygu Bektik, Christothea Herodotou, and Denise Whitelock. "Towards generative AI for course content production: expert reflections." European Journal of Open, Distance and E-Learning (2024): In-press.

[10] Neela, N., T. N. Manjunath, and R. Nishitha. "Adoption of GenAI for Education Domain." In 2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF), pp. 1-6. IEEE, 2025.

[11] Dosaru, Daniel-Florin, Diana-Maria Simion, Andrei-Horia Ignat, Lorina-Cristina Negreanu, and Alexandru-Corneliu Olteanu. "Using GenAI to Assess Design Patterns in Student Written Code." IEEE Transactions on Learning Technologies (2025).

[12] A. Ryan. CIDI Framework ChatGPT Prompt Ultimate Guide for Enhanced Responses, 2024. URL: https://easyaibeginner.com/cidi-framework-chatgpt-prompt-template/