Scaling Academic Decision-Making with NLP: Automating Transfer Credit Evaluations

Nimisha Roy
nroy9@gatech.edu
Georgia Institute of Technology
Atlanta, USA

Omojokun Olufisayo
omojokun@cc.gatech.edu
Georgia Institute of Technology
Atlanta, USA

Huaijin Tu
htu35@gatech.edu
Georgia Institute of Technology
Atlanta, USA

ABSTRACT

Manual processes for evaluating external course syllabi for transfer credit in higher education are time-consuming, inconsistent, and prone to bias. This project leverages Natural Language Processing (NLP) and large language models (LLMs) to automate the transfer credit evaluation process. The system processes external syllabi by embedding course content, conducting similarity searches, and providing structured reasoning for each match. Using techniques such as chain-of-thought reasoning and reflection agents, the system generates similarity scores and detailed explanations to support informed, data-driven decision-making by faculty. Validated against faculty decisions, the system promises to significantly improve the efficiency, consistency, and fairness of transfer credit evaluations. Future directions include expanding the system for advanced standing test evaluations and allowing faculty to query specific course components for more targeted analysis.

ACM Reference Format:
Nimisha Roy, Omojokun Olufisayo, and Huaijin Tu. 2025. Scaling Academic Decision-Making with NLP: Automating Transfer Credit Evaluations. In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 2 (SIGCSE TS 2025), February 26 – March 1, 2025, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3641555.3705245

1 INTRODUCTION AND MOTIVATION

In higher education, manual and subjective processes such as evaluating external course syllabi for transfer credit, reviewing advanced standing test essays, and improving course syllabi by comparing them with similar courses from other institutions present ongoing challenges for faculty, advisors, and domain experts. These tasks are time-consuming and prone to inconsistencies, often leading to inefficient decision-making [3], [6]. NLP and LLMs offer promising solutions to streamline these processes by automating content analysis, identifying gaps, and providing structured outputs to guide academic decisions [4]. While the broad objectives of this project are multifaceted, ranging from improving course syllabi to supporting advanced standing test evaluations, the key focus of this poster is on transfer credit evaluation, an area where automating syllabus comparison can greatly reduce manual workload and improve decision consistency. This is because faculty or academic advisors often manually review transfer student syllabi to determine course equivalency. However, academic advisors, who typically handle these requests, may lack the domain expertise to make final decisions. Meanwhile, faculty members who possess the necessary domain knowledge and decision-making authority are often too busy to efficiently evaluate syllabi in large departments. This process can become labor intensive, inconsistent, and subject to bias [1], as each syllabus must be individually compared to a corresponding university course. Using NLP techniques, such as chain-of-thought reasoning and a reflection agent, this project aims to create a system that automates the transfer credit evaluation process by analyzing the content of external syllabi and providing similarity scores with courses at the home institution. The system also offers detailed reasoning for why a course is or isn't a good match, allowing course coordinators to make informed decisions more quickly and effectively. This improves both the efficiency and fairness of the transfer credit approval process.

2 METHODOLOGY

Our approach to automating transfer credit evaluation is structured around several key phases: data preprocessing, NLP model application, embedding generation, and output generation, as illustrated in Figure1. The system leverages few-shot prompting, chain-of-thought reasoning [7], and reflection agents to assess external course syllabi and compare them to institutional courses using a vector-based similarity search.

2.1 Data Preprocessing

In the preprocessing phase, syllabi in unstructured formats like PDFs are converted to plain text using LangChain's Unstructured PDF loader [2]. The course ID, name, and content are extracted by a reflection agent, ensuring relevant topics, concepts, and skills are included while minimizing hallucinations [5]. Broader objectives and applications are inferred only when they align with the covered topics, capturing course objectives and outcomes coherently.

2.2 NLP Model Application

For NLP model application, the extracted course content is embedded into a 3072-dimensional vector using the text-embedding-3-large model of GPT. These embeddings are stored in the Pinecone Vector Database, creating a searchable vector space. External course syllabi are processed through the same pipeline, generating their embeddings. A cosine similarity search is then conducted between the external course embeddings and the institution's course vectors, retrieving the top three matched courses based on the scores.

SIGCSE TS 2025, February 26–March 1, 2025, Pittsburgh, PA

ACM ISBN 979-8-4007-0532-8/25/2
https://doi.org/10.1145/3641555.3705245

2.3 Chain-of-Thought Reasoning and Output

Once the matches are retrieved, the system applies chain-of-thought reasoning to provide step-by-step explanations for each (external course, internal course) match. This reasoning clarifies why a course is a good or poor match by comparing objectives, learning outcomes, and content depth and ignoring less crucial factors such as textbook differences or course policies. The system produces a GPT-4o structured output in JSON format, detailing a ranked list of (bullet point, explanation) pairs. This output explains the reasons for or against the grant of transfer credit and is visualized in a web-based template for easy comparison by course coordinators, allowing them to make informed decisions more efficiently and consistently.

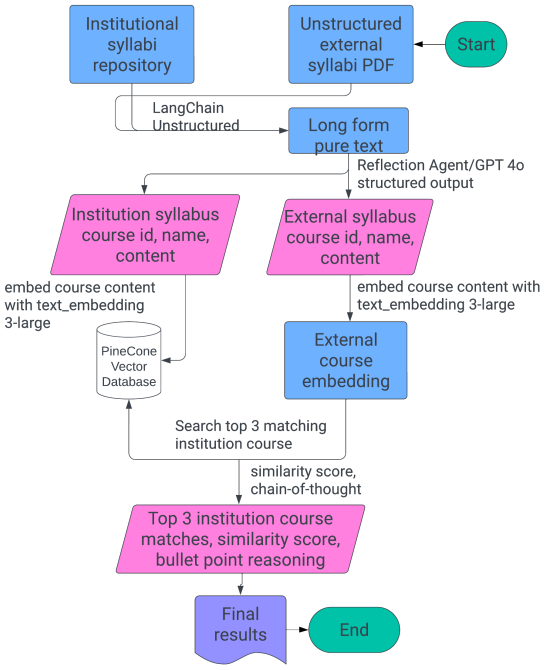

Figure 1: Flowchart of the Transfer Credit Evaluation Process

3 RESULTS AND VALIDATION

The transfer credit evaluation system was tested on a data set of approximately 120 external syllabi and validated against known faculty decisions for five cases. The validation focused on whether the system's top matches and structured reasoning aligned with faculty decisions on transfer credit approval or rejection.

3.1 External Course Evaluation and Reasoning

For the external course "Introduction to Software Engineering" (CS 1111), the system processed the syllabus using the text-embedding-3-large model and compared it with the institution's syllabi. The system identified the following three top matches: CS 2220 - Introduction to Software Engineering with a similarity score of 83.5%, CS 2221 - Objects and Design with 80%, and CS 2222 - Project Implementation with 78.5%. Note that all course IDs are renamed for anonymity. The system provided a detailed reasoning for each match, listing the top reasons for a good and a poor match, ranked in order of importance. To make the technical content accessible to academic advisors without CS expertise, the system also contextualizes each covered or missing topic in terms of its practical significance and learning objectives. CS 2220 was considered a strong match due to its focus on software lifecycle and collaboration in team settings. However, gaps in advanced software architecture and testing methodologies were identified as significant weaknesses. CS 2221 had notable overlap with UML and version control, but the system highlighted the lack of sufficient coverage of object-oriented design principles and patterns as a key area of concern. CS 2222 focused on project management and implementation, but its overall content was less relevant for transfer evaluation.

3.2 Validation Against Faculty Decisions

The system's evaluation was validated against actual faculty decisions. The owners of CS 2220 and CS 2221 courses were approached with the transfer credit request. For CS 2221, the course was rejected for transfer credit based on the system's top reasoning: the external course lacked sufficient emphasis on object-oriented design principles and patterns. CS 2220 course was also rejected, with the top two reasons provided by the system (lack of advanced architecture and limited focus on testing methodologies) aligning with the faculty's decision to reject transfer credit. The accuracy of the system's top reasons for rejection, compared to the faculty's decisions, demonstrates the reliability and effectiveness of the automated evaluation process. The system visualizes the results through a web-based interface to assist course coordinators in reviewing the recommendations. GPT-4o structured output allows the interface to display information clearly for or against transfer credit, supporting fair and efficient decision-making.

Future work will refine prompt engineering to reduce bias in course matching and explore the applicability of this approach beyond CS courses. We also plan to expand features for querying specific course components and explore how machine learning techniques can be applied to assist instructors in improving their syllabi across various domains. In addition, the system will support advanced standing test evaluations.

REFERENCES

[1] D. Chandrasekaran and V. Mago. 2021. Automating Transfer Credit Assessment in Student Mobility–A Natural Language Processing-based Approach. arXiv:2104.01955 [cs.CL] arXiv preprint arXiv:2104.01955.
[2] H. Chase. 2023. LangChain LLM app development framework. Retrieved July, 685(10), p.2023.
[3] L. Liu, R.A. Mendoza, T.R. Martin, and V.M. Miori. 2024. Generative AI-Powered Educational Alignment: A Framework for Matching Syllabus Course Topics with Web Description. In *Proceedings of the 2024 9th International Conference on Distance Education and Learning*. 340–346.
[4] H. Rashtian, A. Hashemi, and L. Macfadyen. 2020. Harnessing natural language processing to support curriculum analysis. In *ICERI2020 Proceedings*. IATED, 1779–1784.
[5] M. Renze and E. Guven. 2024. Self-Reflection in LLM Agents: Effects on Problem-Solving Performance. arXiv:2405.06682 [cs.AI] arXiv preprint arXiv:2405.06682.
[6] B. Ujkani, D. Minkovska, and L. Stoyanova. 2021. Using natural language processing for quality assurance purposes in higher education. In *2021 IV International Conference on High Technology for Sustainable Development (HiTech)*. IEEE, 01–04.
[7] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q.V. Le, and D. Zhou. 2022. Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems*, Vol. 35. 24824–24837.