Enhancing iTextbooks with an Entity Linking Model

Gautam Yadav [1] and Arpit Jain [2]

[1] *Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, 15213, USA*
[2] *PDPM Indian Institute of Information Technology, Jabalpur, 482005, India*

Abstract
As iTextbooks continues to expand, facilitating effective and engaging comprehension for a diverse student base presents a critical challenge. One particular concern lies in balancing course materials to cater to both beginners and advanced learners, with the risk of disengagement and confusion often leading to students skipping or abandoning the video content. To address this issue, we propose an innovative solution utilizing entity-linking machine learning algorithms, like Dexter and GPT-4, within the context of interactive textbooks. We present a system that dynamically links course video or text content with key terms or entities, providing instant access to their definitions in a knowledge base, either Wikipedia or GPT-4, according to an instructor's choice. We have developed a working prototype showcasing an intuitive interface with automatic and customizable entity linking, aiming to make online education more accessible, engaging, and effective. This paper presents the design, architecture, and implementation of this entity linking model, its potential benefits, and future work, including strategies for model evaluation and enhancements.

Keywords
Self-paced learning, Personalized E-learning, Intelligent Textbooks, Entity Linking

1. Introduction

The proliferation of online learning has created new opportunities and challenges in education. Among these challenges is ensuring that course materials are accessible, engaging, and effectively facilitate comprehension for a diverse range of students. The typical structure of online courses often involves video content supplemented by quizzes and exercises. However, maintaining engagement and comprehensibility within these videos is a difficult task. As reported by edX, the average student engagement time with any video is roughly six minutes, irrespective of the video's length [1]. These videos are utilized by a wide array of students, from novices seeking foundational knowledge to experienced learners aiming for course material review [2].

A significant challenge arises when attempting to cater to both beginners and advanced learners within the same video content. Repeated basic information can lead to disengagement for advanced students, while the absence of such information can prove daunting for beginners, hindering their ability to keep pace with the course material [3]. The consequence of this imbalance often manifests in students skipping or abandoning the video content, either due to its difficulty or the time-consuming necessity to refer to external resources for understanding terminology and concepts.

Various strategies have been proposed to enhance video-based learning. Interactive subtitles for notetaking have been suggested [4], along with a video-based learning environment that evaluates student understanding and provides review recommendations [5]. However, these solutions have their

Proceedings Name, Month XX–XX, YYYY, City, Country
EMAIL: gyadav@andrew.cmu.edu (A. 1); arpit.jain@iiitdmj.ac.in (A. 2)
ORCID: 0000−0002−9247−1920 (A. 1); 0000−0003−1101−8479 (A. 2)
© 2020 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

limitations. The former is effective only for revision after understanding the topic, and the latter faces challenges in assessing individual understanding and suggesting appropriate review sections, particularly with large student enrollments.

Previous research on interactive textbooks has shown that personalized and adaptive learning can occur when the platform considers the learner model and content [6]. However, even with these, students often find themselves relying on search engines to clarify terminologies not covered by the interactive systems. This additional process can cause cognitive overload for learners, as they have to shift through numerous non-educational snippets presented by search engine results pages [7]. Our proposed system aims to address this challenge by offering an educational-focused solution, linking terminologies within the content to their definitions in an integrated knowledge base.

Considering these challenges, we propose an innovative solution leveraging entity-linking machine learning algorithms, akin to the Dexter. Dexter is a flexible and configurable entity linking framework that identifies entities in the given text and maps them to their corresponding entries in a knowledge base such as Wikipedia [8]. This involves embedding a dynamic column within the video interface that includes pre-specified terms or entities linked to their corresponding definitions within a knowledge base. This feature enhances interactivity, reduces distraction, and saves time by providing definitions within the same environment, eliminating the need for external resources. The primary drawback of this approach is the potential time and effort required from the course instructor to create the term list for each video. However, the dynamic nature of the entity linking feature mitigates this issue by allowing automatic term list updates in response to video content changes.

This paper focuses on the design and user experience of an interface featuring real-time entity mapping to definitions (using Wikipedia or GPT-4 as a knowledge base) within an online learning platform. The aim is to showcase how this feature can enhance the learning experience, making online education more accessible, engaging, and effective for all students.

2. Architecture

Our proposed system is designed to prioritize the learning experience offered by video lectures. It aims to minimize distractions and maximize understanding by providing in-context definitions to essential terms or entities. Rather than redirecting students to external resources, our system incorporates an interactive transcript that auto-scrolls with the video. Crucially, it highlights and links important entities mentioned in the lecture to their corresponding definitions in a knowledge base.

The recognition and linking of these important entities are achieved by running a background script for each video in the iTextbooks. This process leverages machine learning algorithms, such as the Dexter system for Wikipedia references, or the GPT-4 model for contextual understanding of the entity within the transcript. The integration of these automated linkages within the learning platform not only optimizes the learning experience but also saves the learner's time that would otherwise be spent on searching external sources.

The multi-tier architecture as illustrated in Figure 1 presents a Model Workflow for Entity Linking in Video Subtitles, but it is equally applicable to text instructional content and practice activities. It features three primary layers: the Graphical User Interface (GUI), the backend processing, and the database.

In the GUI layer, instructors can select course videos from the database. The backend processing layer then fetches or generates transcripts for the selected videos, performs entity recognition, generates candidate entities, ranks them, and extracts relevant features from the chosen knowledge base (either Wikipedia or GPT-4, based on the instructor's selection). This information is stored back into the database for quick access.

In the case of text instructional content or practice activities, the system directly proceeds to the entity recognition step after the instructor selects the content, with the remaining steps being identical.

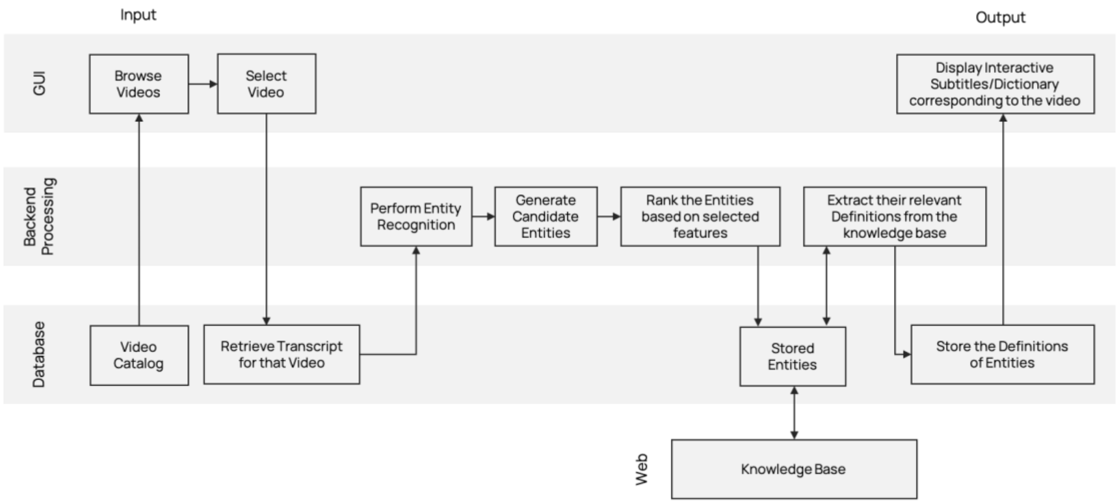

Figure 1: Model Workflow for Entity Linking in Video Subtitles

Our system provides flexibility in choosing the knowledge base. Instructors can set a default option in the system settings (Wikipedia or GPT-4), and they can switch between the two while editing specific entity content in the GUI, ensuring they can choose the most appropriate source based on the given context. This dynamic and comprehensive approach enhances the learning experience by providing contextually relevant definitions directly within the learning platform.

3. System Implementation

To implement this entity linking approach, a web application prototype with two separate interfaces for students and instructors has been developed. The Financial Markets course by Yale University, a highly rated beginner-level course on Coursera, was selected for the implementation. Analyzing reviews of students who rated this course less than 3 out of 5, we observed that many found the concepts unclear and the terminology difficult to grasp. We aim to address this issue by augmenting the course with interactive transcripts that explicitly highlight important entities.

The student interface (see Figure 2) presents the lecture video with an accompanying interactive transcript on the right side. All entities linked to the knowledge base are underlined in the transcript. The learner can click on these underlined entities to view their summarized meanings in a dedicated knowledge section (below the video lecture). This provision allows learners to grasp the contextual definition of complex entities or terms as they are being taught, enhancing focus and comprehension.

In the knowledge section, in addition to the entity's definition, there's a 'more information' button that, by default, directs the student to detailed Wikipedia information about the entity. However, this is not limited to Wikipedia; the button can also redirect students to additional resources specified by the course instructor, such as another module within the iTextbook or a more contextually relevant website.

The instructor interface (see Figure 3) offers a comprehensive overview of each entity visible to the students across all videos in the course. Instructors can add external links to additional resources for extra information. They can also switch between Wikipedia or GPT-4 generated knowledge for each entity, in addition to setting one as a default in their settings.

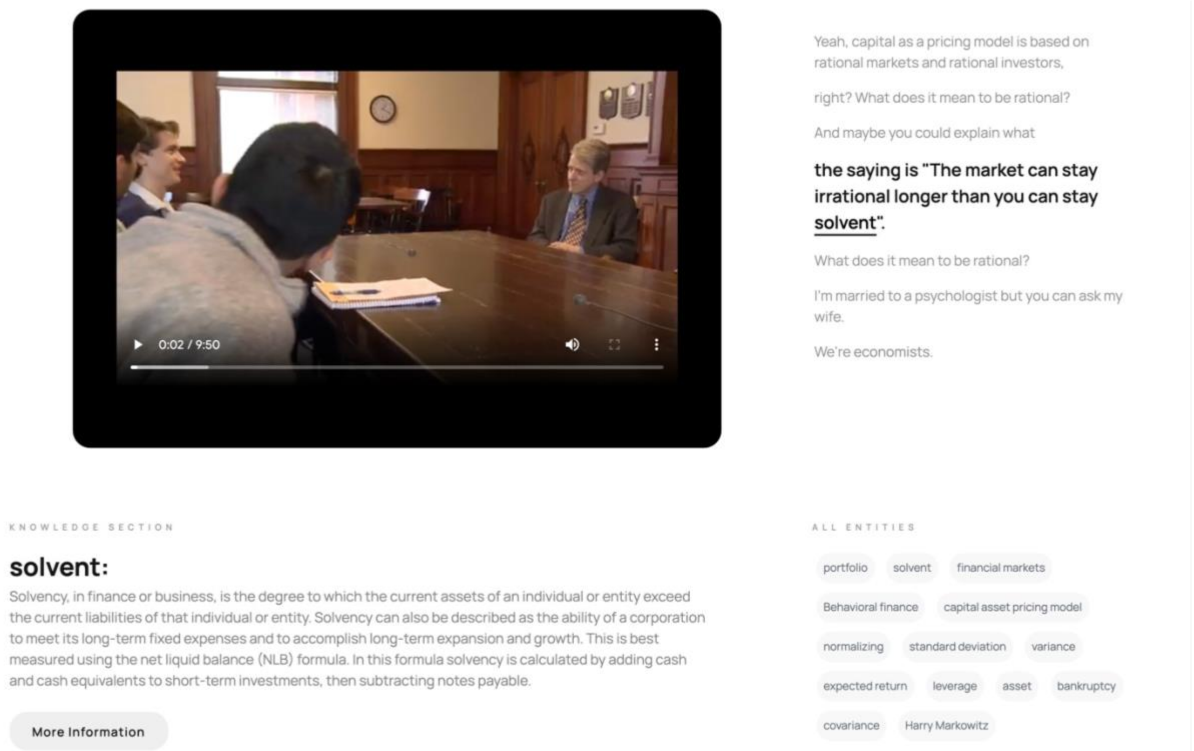

Figure 2: Student Interface showing a lecture video with accompanying interactive transcript. Key entities are underlined and link to their definitions in the knowledge section below the video.

For entity linking, we employed the Dexter system [8], which identifies entities in the transcripts or subtitles and suggests them to the instructor. However, while Dexter excels in entity identification, it occasionally errs in context understanding. For instance, it linked the entity 'solvent' to the Wikipedia page on chemical reactions rather than to financial market solvency. To handle such situations, we added a human layer of validation - the instructor can manually delete or add entities, and correct links as needed.

While Dexter paired with Wikipedia offers a robust and accurate system, GPT-4 offers additional flexibility and comprehensiveness in generating entity descriptions. However, its potential drawbacks, including lower accuracy and credibility in an educational context, led us to offer the instructors a choice between the two. For instance, while both Wikipedia and GPT-4 generated descriptions for the entity 'Harry Markowitz' included details about his birth and prizes won - potentially irrelevant information in the context of the course. By allowing custom entity linking, we aim to save annotation time for instructors and offer more relevance and accuracy in the interactive subtitles.

In this way, our system strikes a balance between the strengths of Dexter/Wikipedia and GPT-4. It leverages the power of machine learning to identify and link entities, while still offering the possibility for human oversight and contextual understanding. This combined approach ensures a comprehensive, yet focused, learning experience for the students.

For entity linking, we used Dexter [8] which identifies entities in the transcripts or subtitles and suggests it to the instructor. Then the instructor has the freedom to manually add the entities that Dexter didn't suggest or delete the trivial entities they don't want. Adding the human (instructor) layer for validation, mistakes due to automated entity recognition are handled. Like in the case of entity called solvent, Dexter linked it to the Wikipedia page referring to chemical reactions rather than the financial market solvency. The instructor can correct this by deleting it manually and linking it to solvency (Wikipedia page for solvent in finance). We found GPT 4 API to be better than Dexter and more comprehensive than Wikipedia pages for generating entity description but due to low accuracy (can make up answers sometimes which appears to be true unlike Wikipedia result so harder to detect and for same prompt often produces different output) and credibility (no evidence for GPT generated instructional content being used) of GPT-4 in teaching context we want to offer instructors a choice whether they want to link using Wikipedia or GPT-4. Adding the feature for custom entity linking here makes it easier for the instructor to manage the interactive subtitles and save the time of annotation for each entity.

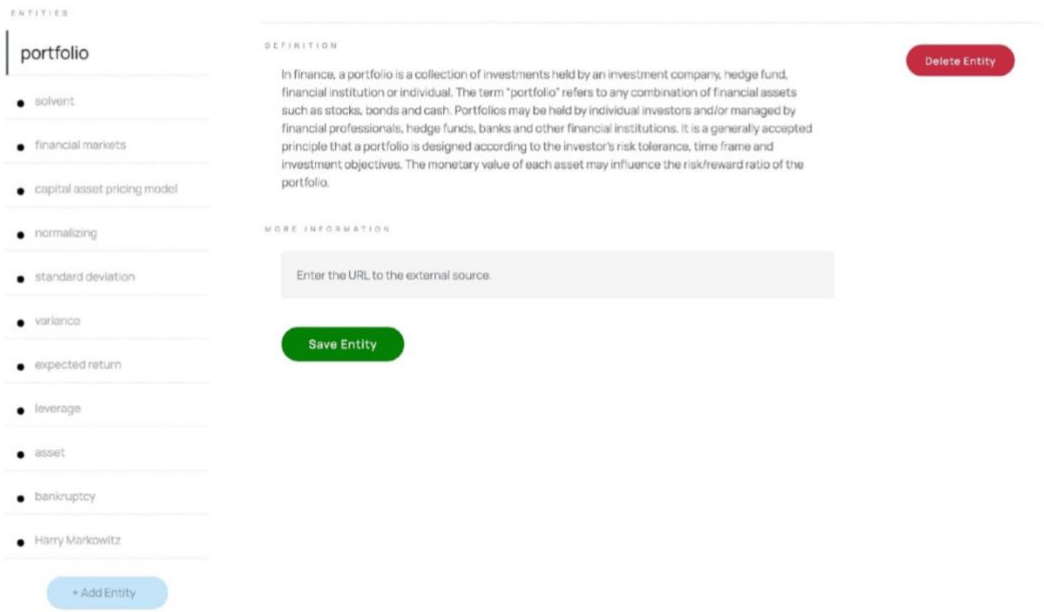

Figure 3: Instructor Interface for managing entity information visible to students across all course videos. The interface allows instructors to add external links, switch between Wikipedia or GPT-4 generated knowledge for each entity and manage entity annotations.

4. Conclusion and Future Work

Our Entity Linking Interface on online learning platforms provides students with an additional resource to comprehend course video and text lecture materials better. When encountering an unfamiliar or complex term, they can quickly access subject-specific or contextual definitions. We presented an intuitive interface that leverages automated entity linking, demonstrated through a semi-automated working prototype. The ultimate goal of this work is to enhance students' learning experience and save their time by offering references within the same learning environment, minimizing distractions. We believe that incorporating this feature into online learning platforms could greatly enhance the role of iTextbooks in future digital transformation trends in education.

We also acknowledge potential limitations of our approach. For instance, while our system is efficient at entity identification, it can falter in understanding context. As an example, Dexter linked the entity 'solvent' to the Wikipedia page on chemical reactions rather than financial market solvency. We added a human layer of validation, allowing instructors to manually delete or add entities and correct links. Additionally, although GPT-4 offers comprehensiveness in generating entity descriptions, its lower accuracy and credibility in an educational context led us to offer the instructors a choice between Wikipedia and GPT-4. As we strive to optimize our system, we aim to reduce the need for manual reviews, improve entity ranking for Wikipedia, and generate better prompts for GPT-4. In future work, we also aspire to explore creating mini iTextbook lessons with practice activities generated by ChatGPT for specific entities to make the instruction active instead of passive.

Looking ahead, we aim to evaluate our system more comprehensively. We intend to collect data on learning gains, time spent reviewing entities, and knowledge base usage patterns to determine the most effective knowledge base strategy for different types of learners and domains. We also plan to measure pre- and post-test gain, time to achieve mastery, and transfer to future learning to assess the impact of our entity linking model on student learning outcomes. Finally, we wish to gather feedback on self-reported engagement, perceived helpfulness of entities, and overall satisfaction to understand if students find the generated content engaging and helpful.

5. References

[1] C. Hazlett, Optimal Video Length for Student Engagement, 2013. URL: https://blog.edx.org/optimal-video-length-student-engagement

[2] Lagerstrom, Larry, Petr Johanes, and Umnouy Ponsukcharoen. "The myth of the six-minute rule: Student engagement with online videos." 2015 ASEE annual conference & exposition. 2015.

[3] Bergmann, Jonathan, and Aaron Sams. Flip your classroom: Reach every student in every class every day. International society for technology in education, 2012.

[4] Deb, Suman, Anindya Pal, and Paritosh Bhattacharya. "Design Considerations for Self Paced Interactive Notes on Video Lectures-A Learner’s Perspective and Enhancements of Learning Outcome." Intelligent Human Computer Interaction: 9th International Conference, IHCI 2017, Evry, France, December 11-13, 2017, Proceedings 9. Springer International Publishing, 2017.

[5] HAYASHI, Yusuke, et al. "Sectional Review Recommendations based on Learner's Comprehension in Video-based Learning."

[6] Ou, Chaohua, Ashok Goel, and David Joyner. "Towards a Pedagogical Framework for Designing and Developing iTextbooks."

[7] Homte, Jaurès Kameni, Bernabé Batchakui, and Roger Nkambou. "Search engines in learning contexts: A literature review." International Journal of Emerging Technologies in Learning (iJET) 17.2 (2022): 254-272.

[8] Trani, Salvatore, et al. "Dexter 2.0: an open-source tool for semantically enriching data." Proceedings of the 2014 International Conference on Posters & Demonstrations Track. Vol. 1272. 2014.