CONREC: Continuous Recommendation of Online Learning Videos Based on Concept Maps

Sungbeom Lee[†], Bumku Choi[†], Jungkyu Han[*] and Sejin Chun[*]

Department of Computer Engineering, DongA University, Busan, South Korea

Abstract

The global proliferation of COVID-19 has catalyzed a substantial transition from traditional educational settings to online learning environments. This shift has precipitated exponential growth in online educational content on video-sharing platforms like YouTube. However, this abundance of content often leaves learners navigating from the massive number of videos. Many learners struggle to identify learning videos that align with their learning objectives. To cope with the challenge, we present CONREC, a prototype application for online learners that recommends the next learning videos based on *concept maps* as a network of knowledge the user has learned. CONREC features an adaptive recommendation that re-ranks the candidates of learning videos based on the combined scores between inter-concepts learned and not learned by the user. We implemented a general-purpose interface that allows learners to continuously watch new learning videos and browser the concept maps of the current state in a visual form.

Keywords

Concept maps, Knowledge graph, Learning video recommendation

1. Introduction

The global spread of COVID-19 has led many students to move from the traditional education hub to the digital education space, ranging from higher education to personal guidance. This has opened new possibilities for Massive Open Online Courses (MOOCs)[1, 2] and Open Educational Resources (OER)[3], innovative educational media that replace the traditional education system[4, 5, 6]. Many learners choose remote learning processes over face-to-face classes, even in the post-pandemic era because of the excellent advantages of online education such as flexibility and convenience.

Recently, a video-sharing platform like YouTube has the potential to provide an elastic and liberal medium for knowledge sharing and instruction[7, 8, 9], where anyone who wants to teach can teach and anyone who wants to learn can access the contents at no or little cost[10]. the video-sharing platforms offer numerous ways in which viewers may engage in learning processes that are self-paced and socially engaging[11]. Many learning videos in the video-sharing platform tend to possess a good range of diversity in terms of topics, formats, and scopes due to the open and social nature of the platform.

However, despite these advantages, learners often find themselves lost in a massive number of educational videos, losing their way toward their learning goals. The massive number of online learning videos can become a hindrance to learning, with the challenge being finding suitable educational materials among countless videos[12, 5, 10]. An online search of learning videos in the video-sharing platforms reveals thousands of videos with varying content, presentation styles, duration, and video quality. It is hard to find learning videos that are suitable to background knowledge, learning objectives, and preferred learning. style[13]. Moreover, navigating through search results usually involves skimming through video content, which means watching parts of the video to determine its appropriateness[14]. This takes a long time for just one video and clearly cannot be extended to a practical scale across multiple videos[15].

Our main contribution are as follows. To address these challenges aforementioned, we propose a CONREC system that recommends the next learning videos based on knowledge(a.k.a *concept*) learned from a number of learning videos. To the best of our knowledge, we are the first to show that CONREC ranks recommended videos based on the user' understood and not-yet-understood knowledge as a concept map, prioritizing videos that address the latter. Secondly, we present a general-purpose pipeline that extracts semantic-enriched, highly-relevant concepts from a learning video. Thirdly, we propose a recommendation method that computes the similarity between the user' concept map and concepts in the learning video. Last, we implement an web-based learning application that supports an adaptive navigation of learning videos as well as a visual interface of concept maps.

Our code and demo of CONREC is available in https://github.com/datascience-labs/conrec.

Fifth Knowledge-aware and Conversational Recommender Systems (KaRS) Workshop @ RecSys 2023, September 18–22 2023, Singapore.

[*]Corresponding author.
[†]These authors contributed equally.
sungbeomlee@donga.ac.kr (S. Lee); choirnony00@gmail.com (B. Choi); jkhan@dau.ac.kr (J. Han); sjchun@dau.ac.kr (S. Chun)
0009-0005-4146-1828 (S. Lee)

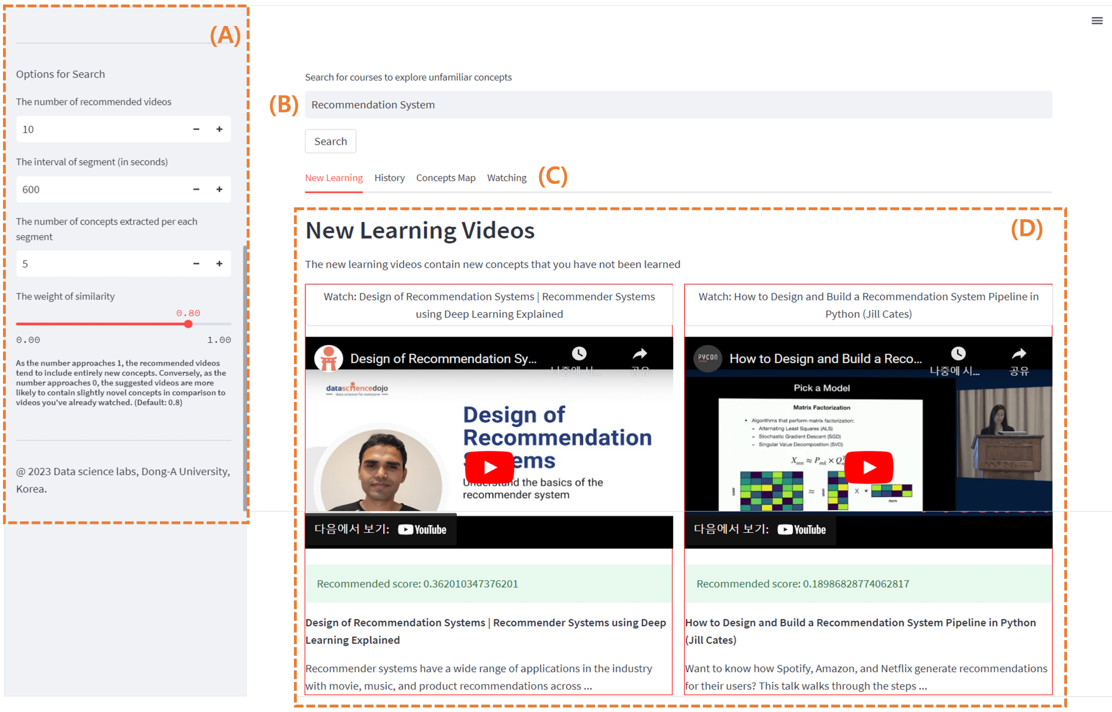

Figure 1: The main interface of the proposed solution. It provides Sidebar (A), Search box (B), Feature tabs (C), and Feature page (D) to learners.

2. CONREC

This section introduces an overview of CONREC. We first describe how to extract semantic and high-relevant concepts within content. The extracted concept can be added to the user' concept map or used to model information of a learning video. Next, we present our recommendation method that computes the similarity between concept vectors as a one-hot representations of concepts. It also helps to re-rank rankings of candidate videos by continuously applying to them.

2.1. Concept Extraction

In this subsection, we model knowledge of the video content as concepts that subjects taught in the learning video. To extract concepts from the video, we introduce data sources used, video segmentation, and concept acquisition.

2.1.1. Data sources

Our data source is YouTube video-sharing platform as our primary source that boasts approximately 37 million pieces of content, encompassing videos in a multitude of languages from across the globe. Notably, YouTube hosts complimentary lectures from globally renowned universities such as MIT, Caltech, and Harvard. Furthermore, YouTube provides an API, which is instrumental in extracting both learning videos and their corresponding videos text.

We extract a limited range of lecture videos to extract videos related to the topic. When a user inputs a search term and presses the search button, the number of videos returned from the YouTube API is retrieved. We selectively filter out videos that are not relevant for learning purposes, specifically those that are either too short (less than 10 minutes) or excessively long (more than 120 minutes). Overly brief videos may not encompass sufficient knowledge content, while exceedingly long videos could potentially impair the concentration of learners.

2.1.2. Video segmentation

Segmentation[5] is utilized to systematically represent the sequential information in learning videos. From the learning videos returned by the search, we extract subtitles, and a boundary-based segmentation is performed on the subtitle text to extract a certain number of concepts. When video texts are provided in the learning video, a list of all texts present in the video is first retrieved. From this list, we determine the existence of video texts and extract the corresponding texts. We have specifically set up fixed interval of 5 minutes, which typically corresponds to approximately 300-400 words. In addition, we have set the margin to include words within a range of 5 seconds in the existing segmentation, preventing them from being included in the following segment. We have made it possible to adjust the range of the margin based on input parameters. The adjustment of the margin can be seen in Figure 1(A).

2.1.3. Concepts acquisition

To obtain concepts from the texts of the segments, we used Wikification[16], a method that utilizes information existing in Wikipedia, such as pages, page texts, anchor texts, and links. It automatically extracts concepts from sentences and identifies semantically similar subjects in Wikipedia based on the words used in the sentences.

2.2. Recommendation Based on Concept Maps

In this subsection, we define and present several concepts related to the concept map-based recommendation method.

A **collection of learning videos** is a finite set of videos related to a specific subject, such as the recommendation system, denoted as $V = \{L_1, L_2, ..., L_n\}$, where L_n represents a learning video. Each video L is considered a document containing its video texts, which are provided in the form of video subtitles or speech scripts.

Concept map represents the set of concepts that can be learned or have been learned from the watched videos. Let us denote the concept map as $K = C \cup C'$, where C and C' indicate sets of learnable or learned concepts, respectively. Both sets contain one or more concepts, referred to as $C = \{c_1, c_2, ..., c_n\}$, where each concept c_n is mentioned in the video texts.

For efficient comparison between concept maps, we vectorize the concept map using a one-hot representation of concepts, which is referred to as the **concept vector** and denoted as $\vec{C}$. The size of $\vec{C}$ is determined by the number of concepts in C.

For clarity, we view **learnable concepts** as concepts that are not clearly understood by a user from he most-recent learning video the user has watched. In other words, learnable concepts are concepts that the user wants to learn from another lecture video. We define learnable concepts from the most-recent learning video they watched as follows.

- Given the most-recent learning video L_r, $\overrightarrow{C_{L_r}}$ is a concept vector in which concepts in the L_r are not yet understood by the user. Otherwise, $\overrightarrow{C'_{L_r}}$ is a concept vector in which concepts in the L_r are understood explicitly by the user.

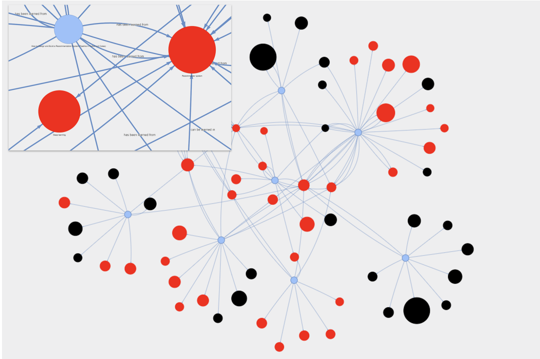

Figure 2: Concept maps: a network of both learnable and learned concepts

Similarly, we present the definition of **learned concepts** as follows.

- $\overrightarrow{C'_{L_n}}$ is the concept vector in which concepts in the learning video L_n understood explicitly by the user.
- $\overrightarrow{C'}$ is the concept vector in which concepts are learned from a multiple number of learning videos the user has watched. Note that $\overrightarrow{C'}$ is all of the concept maps learned from the learning videos that the user watched related to a particular subject.

Given K, L_r, and V, our recommendation method computes the similarity between concept vectors as the following equation.

$$S(K, L_r, V) = \alpha \cdot sim(\overrightarrow{C_{L_r}}, \overrightarrow{C_{L_n}}) + (1-\alpha) \cdot sim(\overrightarrow{C'} - \overrightarrow{C'_{L_r}}, \overrightarrow{C_{L_n}}) \quad (1)$$

, where *sim* function denotes either Jaccard or Cosine similarity computations and α denotes a weight value between 0 and 1.

Here, $sim(\overrightarrow{C_{L_r}}, \overrightarrow{C_{L_n}})$ determines how many learnable concepts are included in the new video. $sim(\overrightarrow{C'} - \overrightarrow{C'_{L_r}}, \overrightarrow{C_{L_n}})$ reduces the importance of previously-learned concepts while maintaining concepts learned from the most-recent learning video.

The weight values can determine the degree of importance of the learning video containing learning concepts. In other words, we can restrict videos containing learnable concepts from being ranked high when recommending the next videos. It is worth noting that we allow the user to adjust α according to the domain.

3. Demonstration

This section presents our implementation.

3.1. Application Overview

CONREC supports an adaptive navigation of video content for learners to explore videos based on the concept map of a user. It searches YouTube videos based on the concept map given by users and return re-rank videos based on the score calculated by our recommendation method. Figure 1 shows four panels of the main interface: (A) User Configuration, (B) Search based on concept map, (C) Feature tabs, and (D) Contents.

3.2. Main Features

In this subsection, we give an explanation of user configuration, search module, and feature tabs.

3.2.1. User configuration

In the sidebar panel, these parameters are used to specify configurations according to the characteristics of the learning domain. There are four options that users can determine (1) the number of videos to search, (2) how many seconds to divide the video to extract key concepts, (3) how many concepts to extract from the sections divided by seconds in the video, and (4) how much weight to give when measuring similarity. Especially, he weight value indicates the learnable concepts in the video. It means value closer to 1 suggests that the video contains learnable concepts, while a value closer to 0 indicates that the video primarily consists of learned concepts. With this parameter, users can tailor and receive video lecture recommendations based on their preferences.

3.2.2. Search module based on concept map

When the user enters a keyword into the search bar and clicks the search button, the system requests the search keyword to the YouTube API. The videos are retrieved according to the number of videos and stored as a video object. Each video object contains the title, URL, video description, and video length. Next, it extracts video texts from the selected videos when they provides subtitles. Otherwise, we obtain it through automatic speech recognition (ASR) provided by the YouTube platform. We then divide the video texts into a fixed number of segments. After the division, it apply the Wikification method to all segments in order to extract concepts depending on the number of concepts. Here, smaller than the number of concepts are extracted when there are not enough number of keywords. The extracted concepts includes a name, URL, and PageRank score.

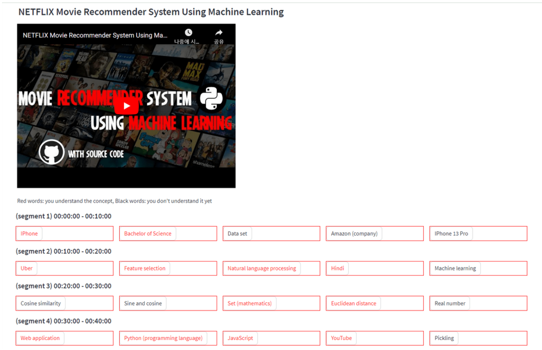

Figure 3: Figure show that Watching tab which Learner could watch learning video where He or She has selected video at New learning tab and could learn the concepts that include in each segments

3.2.3. Feature tabs

The feature tabs consists of four sub-tabs: new learning, history, concept map, and watching tabs.

In the new learning tab, a user can see candidates of learning videos through the recommendation system when searched in the top search bar. In the Figure 1 (D), each video is composed of four components: the watch button, the video, the recommendation score, the video title, and the video description. First, based on the results of a keyword search in the search box, learning videos appear for the learners. When the 'watch button' of a displayed video is clicked, the selected video is classified as a 'watched video', and later, the concepts learned from that video are used for recommendations. Furthermore, the clicked video can be viewed in the 'watching video tab'. The recommended score represents a value calculated by the recommendation algorithm. The higher the recommended score a video has, the higher it is placed.

In the history tab, it displays a list of videos that the user has watched. The list includes the video, video title, a re-watch button, and segment information about the concepts contained in the video. The Re-Watch button allows you to re-watch the video, enabling learners to re-learn from sections they might not have fully grasped initially. Additionally, learners can check segment information for the videos they've watched, which displays information about learnable and learned concepts for each segment of the video. This segment information is provided in table, with each concept marked as either 0 or 1, depending on whether it was understood by the learner.

In the concept map tab, it visualizes the concepts from the watched videos as a network graph. In Figure 2, Each sky blue node represents the videos that have been watched, each black node represents the concepts that were not understood from the video, and red nodes signify the concepts that were understood. The size of these concept nodes increases based on the PageRank value obtained through Wikifier, with higher values resulting in larger nodes. If different videos share the same concept and link to the same Wikipedia concept, they are interconnected.

In the watching tab, it allows users to view the video they selected from either the new learning tab or the history tab. In the watching tab, there is the selected video, segments related to the video, and the concepts contained within each segment. The number of segments and concepts displayed reflect the settings made in the options of Figure 1(A). Each segment contains concepts up to the set number, and they are displayed in order of importance. Initially, all concepts are in a 'learnable' state and are shown in black. As users watch the video and understand a concept, they can click on it, turning it red, which signifies a 'learned concept'. The concepts learned in the Watching tab are reflected in the history tab and the concept map tab. These learned concepts are then used as a basis for recommending new learning videos

4. Learning process based on concept maps

In this section, we describe the learning process of online videos based on concept maps. The following user scenario illustrates a continuous learning process.

In the video-sharing platform, a user searches for learning videos relating to keywords of the concept she wants to learn (e.g., recommender system). The video-sharing platform retrieves multiple videos. She can select the first video from them to learn about the concept and watch the selected video. As shown in Figure 3, she can interactively mark either learnable or learned concepts in the middle or at the end of the video she is watching. Her concept map is incrementally created whenever she marks either learned or learnable concepts.

Regardless of whether she watches the video completely or not, she can select the next videos to understand learnable or learned concepts. At that time, the rank of the lecture videos was rearranged according to the current state of her concept map. Here, the user settings aforementioned affect the rank of the videos.

The visualized concept map allows her to check what she has learned. In addition, the co-occurrence of concepts in the watched videos is displayed. Moreover, she can watch the learning video again in her watched history if she either did not understand some concepts or re-learn concepts. She searches for keywords of either learnable or learned concepts to obtain additional candidates for lecture videos.

Last, she can perform these learning processes repeatedly until she understands the concept. The learning process terminates when concepts are completely understood by either the system or her intention.

5. Conclusion and future work

To summarize, we have proposed CONREC, a recommendation system tailored for exploring learning videos based on concept maps. First, CONREC has identified suitable learning videos based on knowledge the user has learned but also re-ranked them considering his understanding. Second, CONREC has navigated video content adaptively, dividing videos into segments and visualizing concepts as network graphs. Third, we have defined both learnable and learned concepts and represented the concept vector as a one-hot representation of concepts. Last, our graphical interface has ensured users can customize their learning experiences, enhancing content relevance, and engagement.

In the future work, we will improve CONREC that aligns video recommendations with potential career paths. Our future work provides learners a personalized learning path that not only enriches their knowledge but also enhances their career prospects.

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (No.2021R1F1A1050937)

References

[1] L. Pan, C. Li, J. Li, J. Tang, Prerequisite relation learning for concepts in moocs, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2017, pp. 1447–1456.

[2] J. Yu, Y. Wang, Q. Zhong, G. Luo, Y. Mao, K. Sun, W. Feng, W. Xu, S. Cao, K. Zeng, et al., Mooccubex: a large knowledge-centered repository for adaptive learning in moocs, in: Proceedings of the 30th ACM International Conference on Information & Knowledge Management, 2021, pp. 4643–4652.

[3] S. Roy, M. Madhyastha, S. Lawrence, V. Rajan, Inferring concept prerequisite relations from online educational resources, in: Proceedings of the AAAI conference on artificial intelligence, volume 33, 2019, pp. 9589–9594.

[4] C. Impey, M. Formanek, Moocs and 100 days of covid: Enrollment surges in massive open online as-

tronomy classes during the coronavirus pandemic, Social Sciences Humanities Open 4 (2021) 100177.
[5] M. Pérez Ortiz, S. Bulathwela, C. Dormann, M. Verma, S. Kreitmayer, R. Noss, J. Shawe-Taylor, Y. Rogers, E. Yilmaz, Watch less and uncover more: Could navigation tools help users search and explore videos?, in: Proceedings of the 2022 Conference on Human Information Interaction and Retrieval, 2022, pp. 90–101.
[6] J. Gong, Y. Wan, Y. Liu, X. Li, Y. Zhao, C. Wang, Y. Lin, X. Fang, W. Feng, J. Zhang, et al., Reinforced moocs concept recommendation in heterogeneous information networks, ACM Transactions on the Web 17 (2023) 1–27.
[7] Y. Chtouki, H. Harroud, M. Khalidi, S. Bennani, The impact of youtube videos on the student's learning, in: 2012 international conference on information technology based higher education and training (ITHET), IEEE, 2012, pp. 1–4.
[8] T. Jones, K. Cuthrell, Youtube: Educational potentials and pitfalls, Computers in the Schools 28 (2011) 75–85.
[9] C. Snelson, The benefits and challenges of youtube as an educational resource, The Routledge companion to media education, copyright, and fair use 48 (2018) 109–126.
[10] C.-L. Tang, J. Liao, H.-C. Wang, C.-Y. Sung, W.-C. Lin, Conceptguide: Supporting online video learning with concept map-based recommendation of learning path, in: Proceedings of the Web Conference 2021, 2021, pp. 2757–2768.
[11] P. G. Lange, Informal Learning on YouTube, John Wiley Sons, Ltd, 2018, pp. 1–11.
[12] M. Schwab, H. Strobelt, J. Tompkin, C. Fredericks, C. Huff, D. Higgins, A. Strezhnev, M. Komisarchik, G. King, H. Pfister, booc. io: An education system with hierarchical concept maps and dynamic non-linear learning plans, IEEE transactions on visualization and computer graphics 23 (2016) 571–580.
[13] D. Singh, A. Gupta, C. Jawahar, M. Tapaswi, Unsupervised audio-visual lecture segmentation, in: 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), IEEE, 2023, pp. 5221–5230.
[14] B. Zhao, S. Lin, X. Luo, S. Xu, R. Wang, A novel system for visual navigation of educational videos using multimodal cues, in: Proceedings of the 25th ACM international conference on Multimedia, 2017, pp. 1680–1688.
[15] D. Mahapatra, R. Mariappan, V. Rajan, K. Yadav, S. Roy, Videoken: Automatic video summarization and course curation to support learning, in: Companion Proceedings of the The Web Conference 2018, 2018, pp. 239–242.
[16] J. Brank, G. Leban, M. Grobelnik, Annotating documents with relevant wikipedia concepts, Proceedings of SiKDD 472 (2017).