Search and Recommendation Systems with Metadata Extensions

Woo-Hyeon Kim*, Joo-Chang Kim**

Division of AI Computer Science and Computer Engineering, Kyonggi University, South Korea
*** Contents Convergence Software Research Institute, Kyonggi University, South Korea*

whkim712@kyonggi.ac.kr, kjc2232@naver.com

Abstract— **This paper proposes an AI-based video metadata extension model to overcome the limitations of video search and recommendation systems in the multimedia industry. Current video search and recommendation utilize pre-added metadata. Metadata includes filenames, keywords, tags, genres, etc. This makes it impossible to make direct predictions about the content of a video without pre-added metadata. These platforms also analyze your previous search history, viewing history, etc. to understand your interests in order to serve you personalized videos. This may not reflect the actual content and may raise privacy concerns. In addition, recommendation systems suffer from a cold start problem, which is the lack of an initial target, as well as a bubble effect. Therefore, this study proposes a search and recommendation system by expanding metadata in videos using techniques such as shot boundary detection, speech recognition, and text mining. The proposed method selects the main objects required by the recommendation system based on the object frequency and extracts the corresponding objects from the video frame by frame. In addition, we extract the speech from the video separately, convert the speech to text to extract the script, and apply text mining techniques to the extracted script to quantify it. Then, we synchronize the object frequency and the transcript to create a single contextual data. After that, we group videos and clips based on the contextual data and index them. Finally, we utilize Shot Boundary Detection to segment videos based on their content. To ensure that the generated contextual data is appropriate for the video, the proposed model compares the extracted script with the video's subtitle data to check and calibrate its accuracy. The model can then be fine-tuned by tuning and cross-validating the hyper-parameter to improve its performance. These models can be incorporated into a variety of content discovery and recommendation platforms. By using expanded metadata to provide results close to a search query and recommend videos with similar content based on the video, it solves problems with traditional search, recommendation, and censorship schemes, allowing users to explore more similar videos and clips.**

Keyword—**Recommendation, Speech Recognition, Contextualized Data, Metadata**

Woo-Hyeon Kim () was born in Geoje-si, Gyeongsangnam-do, South Korea, July 12, 2003.
In 2022, he enrolled in the Division of AI Computer Science and Computer Engineering at Kyonggi University, majoring in Computer science.
His educational background includes:
      - Bachelor's Degree:Computer Science, Kyonggi University, South Korea, 2022
His Research interests encompass Data Mining, Big Data, and Anomaly Detection

Joo-Chang Kim has received B.S. and M.S. degrees from the School of Computer Information Engineering, Sangji University, South Korea in 2014 and 2016, respectively. He has received Ph.D. from Department of Computer Science, Kyonggi University, South Korea in 2021. He is currently a research professor in Contents Convergence Software Research Institute, Kyonggi University. Since 2021, he is currently a lecturer in the Department of Software Convergence Engineering, Inha University, South Korea. His research interests include data mining, data management, knowledge systems, machine learning, deep learning, big data, healthcare, and recommendation systems.