May 2023

LLM-based Comment Summarization and Topic Matching for Videos

Mason Henry Dimarco

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

LLM-based Comment Summarization and Topic Matching for Videos

ABSTRACT

Large language models (LLMs) have shown substantial improvements on text summarization and classification benchmarks over prior techniques. LLMs can be used for topic detection for online video which can help improve video recommendations. This disclosure describes techniques to parse comments associated with a particular video in addition to video topic detection. The automated parsing enables acquisition of auxiliary information about the context in which the video is made available. The topics/keywords can be complementary to or can extend upon topics detected from the video and/or video metadata. A combination of the video topics and comment topics can be utilized for video recommendations and can identify potential audiences that may not otherwise be matched to a video.

KEYWORDS

- Text summarization

- Topic detection

- Natural language processing

- Topic extraction

- Large language model (LLM)

- Video topic

- Comment topic

- Transformer model

BACKGROUND

Text summarization and topic matching are widely explored areas of natural language processing (NLP) research. These areas were difficult for general purpose solutions. However, the rise of transformer-based large language models (LLMs), with their vast scale, enables rich automated understanding of natural language text. Newer and larger models have shown substantial improvements on text summarization and classification benchmarks. LLMs are already used to generate transcripts of videos online, to enable better automated understanding of the topics covered in the videos. Such automated detection of topics in a video can help improve video recommendation (what videos are to be shown next to a viewer) and is also useful to determine the preferences of the viewer.

DESCRIPTION

This disclosure describes techniques to parse comments associated with a particular video in addition to video topic detection. The automated parsing enables acquisition of auxiliary information about the context in which the video is made available. The topics/keywords can be complementary to or can extend upon topics detected from the video and/or video metadata such as video title, video description, video thumbnail, etc. A combination of the video topics and comment topics can be utilized for video recommendations and can identify potential audiences that may not otherwise be matched to a video.
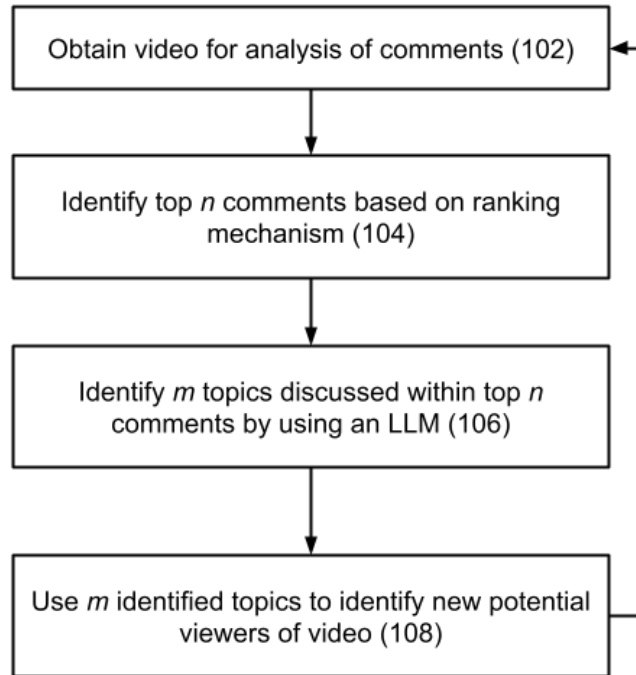
Fig. 1: LLM-based comment summarization and topic matching

Fig. 1 illustrates an example technique of LLM-based comment summarization and topic matching. The described technique can be used to identify potential new viewers for a video. A video and associated comments (e.g., on a webpage where the video is hosted or embedded) are obtained (102). Using a suitable ranking mechanism, the top n comments are identified (104). The top n comments are parsed using a suitable LLM to identify m topics within the top n comments (106). These m topics are used to identify potential new viewers for the video (108). For example, with user permission, the topics can be matched against user interests for users that have not viewed the video. The techniques separate comments that are semantically similar to the topics in the video or associated metadata from comments that are semantically different. The identification of comments that are semantically different enables enlarging the set of recommendations and ensures topic diversity (making sure that the top comments are not all discussing the same narrow topic).
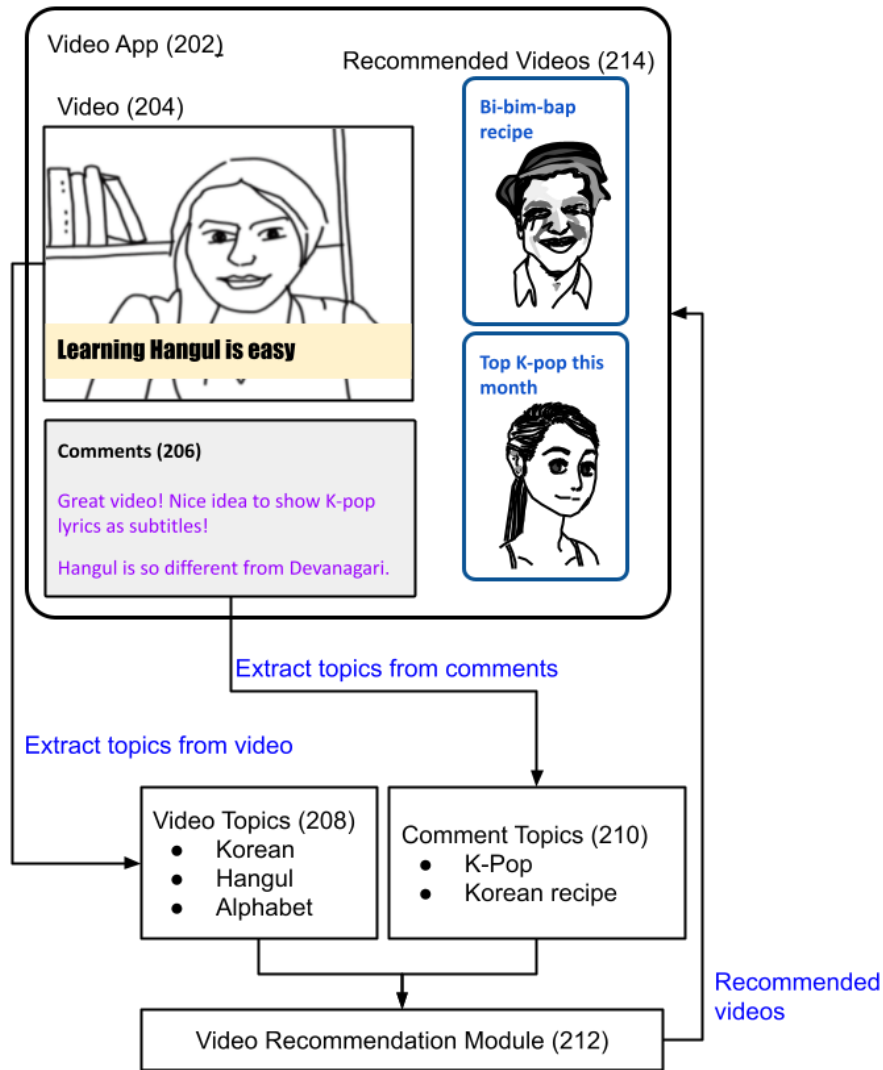
Fig. 2: Video recommendation based on topic identification from video and comments

Fig. 2 illustrates an example of the topics identified from comments associated with a video being utilized to recommend the video to potential new viewers. Within a video app (202), a user watches a particular video (204). In the example of Fig. 2, the video is about how learning Hangul, the South Korean alphabet, is easy. Having watched this video, both the current viewer and other viewers may leave comments on the video (206). The transcript of the video is analyzed using an LLM to identify topics within the video (208). For example, the LLM might

detect that the video is about learning the Korean language, the Hangul alphabet, and about alphabets in general.

When the comments from the video are analyzed to extract topics (210), topics such as K-pop, or about Korean recipes are identified. In this case, the topics of the comments are distinct from those of the video or video metadata. The topics identified from the video and the comments are provided to a video recommendation module (212) which identifies one or more videos that are to be shown (214) to the user.

CONCLUSION

This disclosure describes techniques to parse comments associated with a particular video in addition to video topic detection, e.g., using a LLM. The automated parsing enables acquisition of auxiliary information about the context in which the video is made available. The topics/keywords can be complementary to or can extend upon topics detected from the video and/or video metadata. A combination of the video topics and comment topics can be utilized for video recommendations and can identify potential audiences that may not otherwise be matched to a video.

REFERENCES

1. Serbanoiu, Andrei, and Traian Rebedea. "Relevance-based ranking of video comments on YouTube." In *2013 19th International Conference on Control Systems and Computer Science,* pp. 225-231. IEEE, 2013.

2. Culbertson, Gabe, Wei Peng, and Nicolas Crowell. "Video anchors." U.S. Patent Application No. 17/069,638 filed Oct 13, 2020.