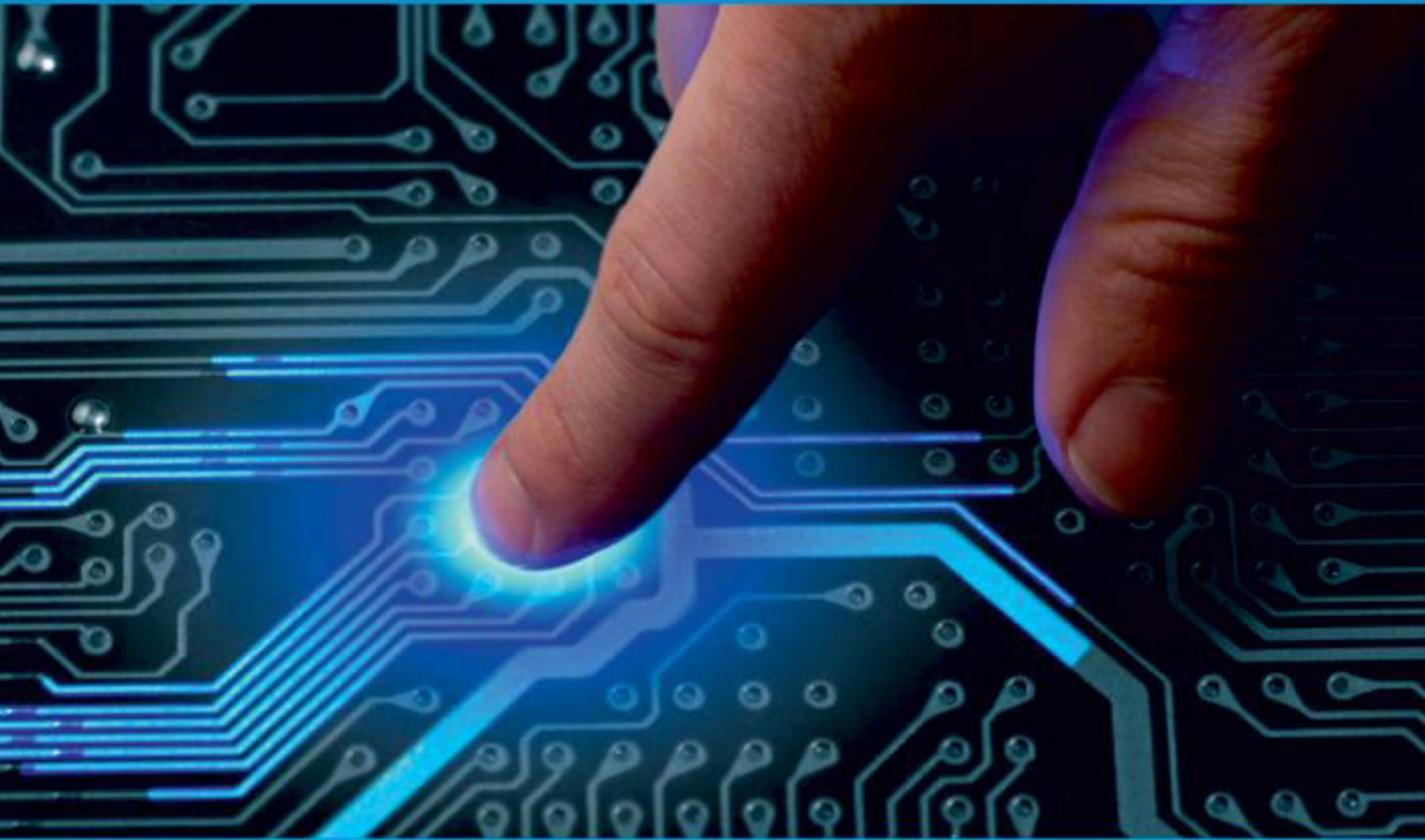
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907213 |

Lecture Video Summarization using Bidirectional Encoder Representations from Transformers Extractive Summarizer

Raghul Adhitya PJ, Rohan P, Sai Suresh MV

Student, Department of Computer Science and Engineering, Velammal Engineering College , Surapet, Chennai, Tamil Nadu, India

Student, Department of Computer Science and Engineering, Velammal Engineering College , Surapet, Chennai, Tamil Nadu, India

Student, Department of Computer Science and Engineering, Velammal Engineering College , Surapet, Chennai, Tamil Nadu, India

ABSTRACT: Watching videos for our own purposes to learn new things has become vital in our day to day lives to learn faster. However, to search according to our preferences mainly when searching lectures or short videos to learn concepts faster, it takes time to scour many videos to find appropriate to one user needs. To speed up the searching process, we are proposing a system to briefly summarise by implementing Lecture video summarization to summarize the contents said in a video and provide a two-line description, video accuracy, and a brief summary. By the information provided by our algorithm, it will be a quick way to search video. The proposed system is implemented by the BERTsummarizer algorithm to process large text information for extracting the necessary information and the Maximum-Cosine-Value algorithm (our own algorithm) for producing cosine values to give rating for the video. This information will be a key factor to browse videos faster.

KEYWORDS: NLP, BERT, Maximum-Cosine-Value, CBOW, YAKE

I. INTRODUCTION

Educational videos are a plentiful supply thanks to the increase of Massively Online Open Courses and university uploads. These resources include the promise of democratizing education by ensuring that distance and economy do not prevent a path to quality content. It is there that breaking down a lecture into topically identified segments enhances watcher engagement. Moreover, researchers have discovered that students visit multiple times to certain periods in the lecture video to reexamine concepts. Therefore, there is a requirement to summarize the lecture video, so that a student can promptly decide which lecture to view and navigate to specific content. Extracting and abstracting data within these lectures into a short form could also improve existing search engines in future. Data extraction from lecture videos is a challenging problem. Many lectures are self-recorded and do not have general production and transcript annotation and often involve a single fixed camera covering the whiteboard. Hence, there is a demand to summarize lecture videos either through transcripts or through audio recognition. In this research, we explored a methodology and built a framework to extract summaries from lecture videos.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907213 |

II. LITERATURE SURVEY

Analysis of Supervised Regression *Nick Su and Ismael Menjiva*, December 8, 2014, results showed that linear regression and SVR find correlations in the data, but are afflicted by the amount of noise in our features and labelling and noise issues in data are generally difficult to remove with linear regression algorithms. Numerous smaller machine learning projects could be conducted to reduce noise in the final regression model.

Text Summarization Techniques: A Brief Survey *Mehdi Allahyari, Mehdi Assef, Saeid Safaei, Krys Kochut, Elizabeth D.Trippe*, July 28, 2017, It is difficult for humans to summarize large amounts of text. Thus, there is an immense need for automatic summarization tools in this age of information overload.

A Survey on Recognizing Textual Entailment as an NLP Evaluation *Adam Poliak*, October 6, 2020, Future RTE datasets targeting specific phenomena that contain scalar RTE labels from multiple annotators can provide even more insight into contemporary NLP model

Abstractive Summarization of Spoken and Written Instructions with BERT *Alexandra Savelieva, Bryan Au-Yeung, Vasanth Ramani,* August 26, 2020, The quality of the test output is comparable to YouTube summaries. Human authors are prone to making language use errors. The advantage of using abstractive summarization models allows mitigating some issues with the video author's grammar.

III. SYSTEM STUDY

Tokenization is the process of breaking text documents apart into those pieces. In-text analytics, tokens are most frequently just words. A sentence of 10 words, then, would contain 10 tokens. For deeper analytics, however, it's often useful to expand your definition of a token. Tokens can be Words, Punctuation, Hyperlinks, Possessive markers Tokenization is language-specific, and each language has its own tokenization requirements.

Sentence boundary disambiguation (SBD), also known as sentence breaking, sentence boundary detection, and sentence segmentation, is the problem in natural language processing of deciding where sentences begin and end. Natural language processing tools often require their input to be divided into sentences. However, sentence boundary identification can be challenging due to the potential ambiguity of punctuation marks. In a paragraph content, processing (i.e) formatting the contents and uniforming the content is important in text processing, in order to achieve it. The following steps must be done: Remove whitespaces, Remove HTML tags, Convert accented characters, Remove special characters, Remove words, Remove stop words.

Text analytics techniques are helpful in analyzing sentiment at the entity, concept, or topic level and in distinguishing opinion holders and opinion objects. Quantitative text analysis is a set of techniques stemming from the social sciences where either a human judge or a computer extracts semantic or grammatical relationships between words in order to find out the meaning or stylistic patterns of, usually a casual personal text for the purpose of psychological profiling etc.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907213 |

IV. PROPOSED SYSTEM

MODULES

A) Retrieval of YouTube transcription
B) Extractive Summarizer
C) User Interface

MODULE DESCRIPTION: A module is a separate unit of software or hardware. Characteristics of modular components include portability and interoperability which allows them to function in another system with the components of other systems.

Retrieval of YouTube transcription: YouTube transcriptions are retrieved using the transcript API, the transcript holds the text spoke and time interval for each sentence, these transcripts are parsed in a well-formatted manner to be sent to the extractive summarizer. The title of the YouTube video is also extracted in this module, which will be processed in the further modules.

Extractive Summarizer: The Extractive Summarizer system is a processing module that summarizes the retrieved YouTube transcription. Extracting the summary means identifying a particular sentence from a corpus of sentences on the transcription. The text extracted from the module will be processed and well-punctuated to be passed into the model. The punctuations are carried out by a recurrent neural network whereas the summary is extracted using the BERT Extractive Summarizer. The summary extracted will be evaluated based on certain evaluation matrices with the title of the YouTube video to provide a context reliability score.

User Interface: This is a presentation module that encapsulates the application of the product, users can provide a link to their required YouTube video or upload a video to which they want to get a summary. The request is handled by a flask hosted website, which redirects the request to the processing modules. The user interface is designed in a manner to provide the best user experience and interaction, using keyword navigation and precise summary.

V. SYSTEM ARCHITECTURE

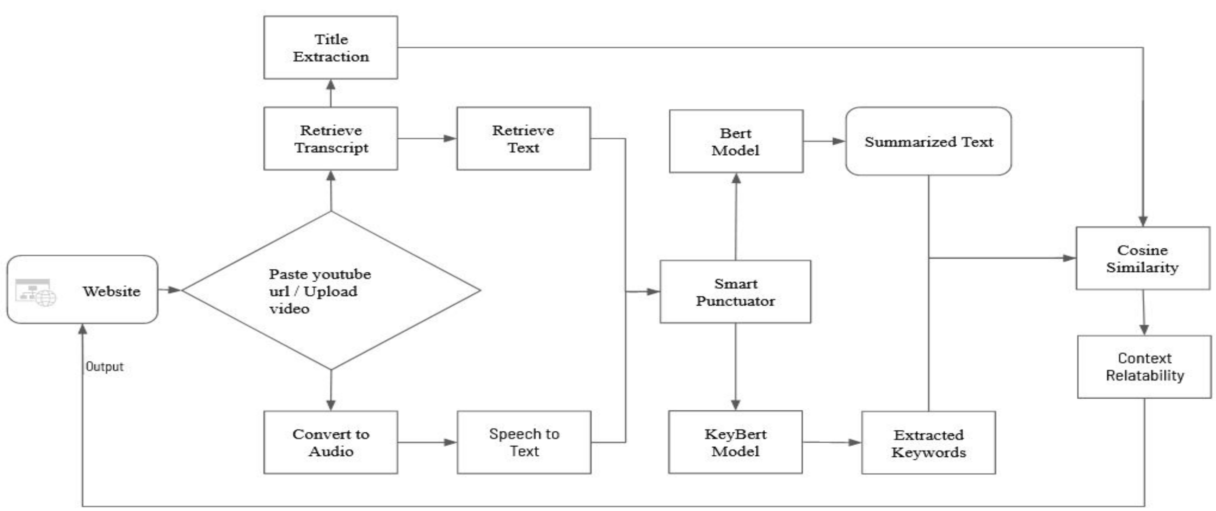


Fig1.0

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907213 |

VI. OUTPUT

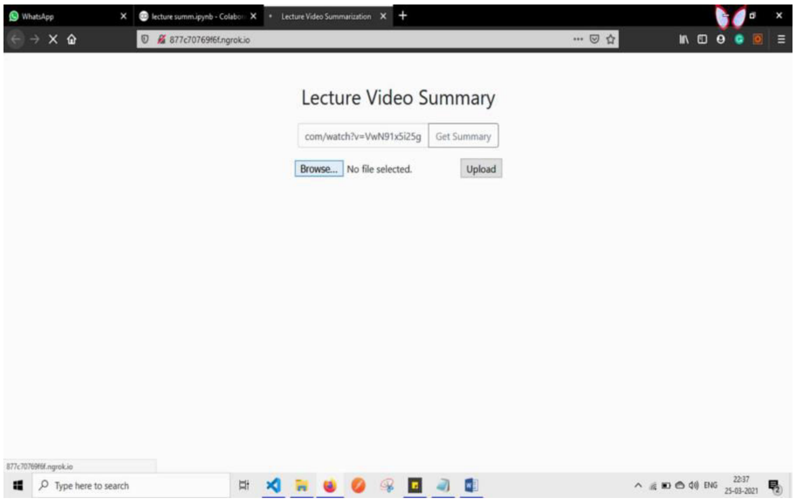


Fig 1.1

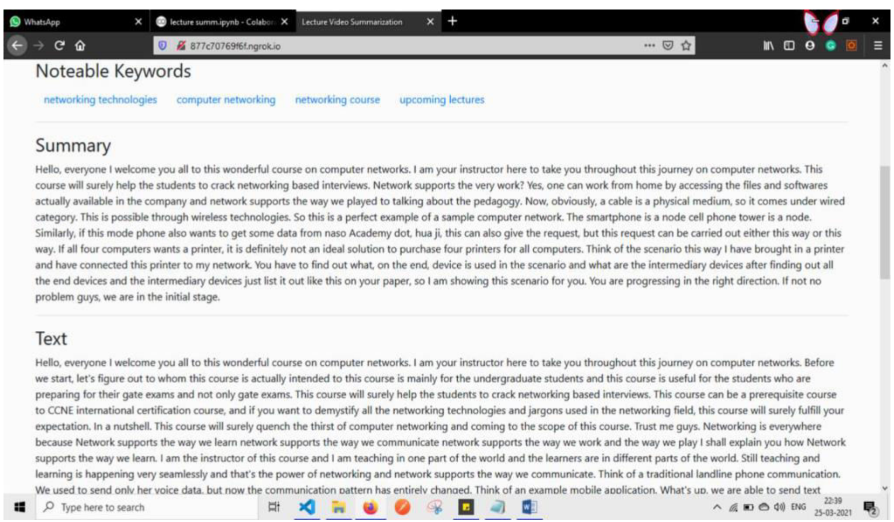


Fig 1.2

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 7, July 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0907213 |

Fig 1.3

VII. CONCLUSION AND FUTURE WORKS

Therefore, the lecture video summarization is successfully implemented using YouTube api transcriptions, punctuations, BERT extractive summarizer, KeyBERT Extractor, Cosine Similarity, Flask Framework, and other UI Components. OCR (Optical Character Recognition) can be implemented in the videos to extract top keywords for better summarization. For example, handwritten characters, words/characters present in the lecture video presentation. This extraction of keywords can be used as meta-data for the lecture video for any search engine. The database can act as a central repository for storing the metadata (i.e) the extracted keywords.

REFERENCES

[1] Alexandra Savelieva, Bryan Au-Yeung, Vasanth Ramani (2020), “Abstractive Summarization of Spoken and Written Instructions with BERT”.
[2] Adam Poliak (2020), “A Survey on Recognizing Textual Entailment as an NLP Evaluation”.
[3] Anish Mishra, Pushpak Bhattacharyya (2017), “Deep learning Techniques in Textual Entailment”.
[4] Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova (2018), “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”.
[5] Mehdi Allahyari, Seyedamin Pouriyeh, Mehdi Assefi (2017), ”Text Summarization Techniques: A Brief Survey”.
[6] Nick Su, Ismael Menjivar (2014), “Predicting Lecture Video Complexity: Analysis of Supervised Regression”.
[7] Vasile Rus, Art Graesser and Philip M. McCarthy and King-Ip Lin (2005). “A Study on Textual Entailment”.
[8] Vivian S. Silvaa, Andr´e Freitasb and Siegfried Handschuh (2018) ,”XTE: Explainable Text Entailment”.

Impact Factor: 7.542

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details