



## AUTOMATIC KEYPHRASE EXTRACTION

Divya Krishna<sup>1</sup>, Krithika S Shetty<sup>2</sup>, Pooja L Devadiga<sup>3</sup>, Sushmitha U<sup>4</sup>  
Shri Madhwa Vadiraja Institute of Technology and Management, Udupi, India.

### Abstract

**Keyword extraction is the most important chore while working with video. By extracting keywords, it benefits the viewers to know the main content of video while watching it. The technique is based on regular expression grammar rule approach to recognize the chunks of noun in the video. Extracting the keywords help in summarizing the video and helping us understand video in brief. Platform used for key phrase extraction system is python and java programming language. The transcripts will be available in the file. A framework for key expression extraction is structured that utilizes a regulated AI calculation.**

**Keywords -FFMPEG library, CTC, Deepspeech, DNN, NLTK library, spaCy feature extractor**

### I. INTRODUCTION

With the unpredictable development of online data, it in some cases gets easy to choose where to go to discover data. It is infeasible for clients to peruse each report of the outcome set to decide if they may be helpful. Catchphrase extraction is entrusted with the programmed distinguishing proof of terms that best depict the substance of the video. Keyphrases give a short depiction of a report's substance.

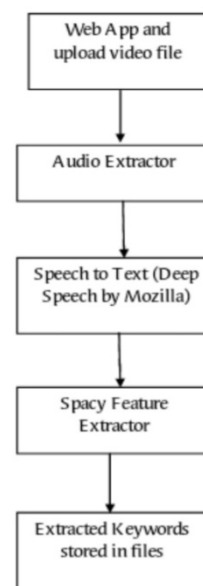
In this proposed venture of programmed keyphrase extraction, is a principal literary data handling task that removes keyphrases consequently from video with a target that these representative expressions will survey its substance.

To improve the talk program viability to understudies, video addresses are frequently joined by address notes, slides or prospectus books. A couple of cutting-edge address

programs arrange text with address video and furthermore permit search inside the content. Be that as it may, these frameworks are of little use if for example, an understudy needs to rapidly check the substance of a specific talk in a progression of talks.

Key phrases are normally chosen in specialized records, for example, diary articles and specialized reports, to feature key subjects and ideas. So also, when applied to addresses, keyphrases can give significant data and can permit understudies to recognize addresses that compare to their learning prerequisites. Mechanizing the procedure of keyphrase extraction in address records presents a few difficulties. The absence of legitimate substance organizing, defects in language use, casual nature and varieties of style across spaces are basic highlights of talks that hinder execution of data recovery frameworks.

### II. METHODOLOGY



**Fig 1: Block diagram of proposed methodology**

As shown in fig 1, the proposed system consists of audio extractor, speech to text convertor, spacy feature extractor and extracted keywords are the stored in separate file.

Video is given as input file to audio extractor. To extract audio from video, we use FFMPEG library. FFMPEG is multimedia framework used to process and convert video and audio, capable of taking input from live source. It can also convert arbitrary sample rates and resize video.

Extracted audio file is fed as input to speech to text converter. Mozilla deep speech is used to convert audio into text file. Deep speech is an Automatic Speech recognition engine which uses a bidirectional RNN implemented with tensor flow. Converted text file is fed as input to spaCy feature extractor. It extracts nouns, adjectives from the text file using NLP and then stored in separate file.

**Pre processing:** In preprocessing raw audio waveform signal is taken as input and is converted it into a log-spectrogram of N-timesteps and N-frequency-features. N-frequency-features can be allotted in the model's design record as "num-sound highlights" boundary, N-time steps relies upon unique sound document's term.

**Spectrogram:** The visual representation of the spectrum of frequencies of a signal (audio input) as it varies with time is called spectrogram. It is general representation of an audio signal.

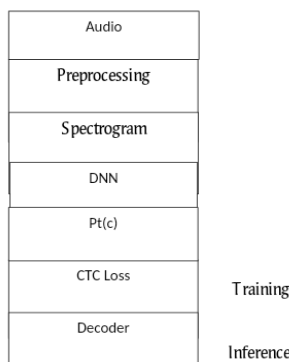


Fig 2: Deep speech

**Evaluation metrics:**To check the model performance Evaluation metrics are used. WER (Word Error Rate) is used as evaluation metric. Decoder changes over a likelihood dispersion over characters into text.

**DNN:** Basic working of convolution layer (CNN) is like neural networks, which are comprised of neurons with learnable loads and inclinations. The Deep Neural Network (DNN) part produces a probability distribution  $P_t(c)$  over vocabulary characters per each time step.

**CTC:** Greedy decoder and Beam search decoder are the 2 of decoders that are generally utilized with CTC-based models.

**Greedy decoder:** An greedy decoder yields the most likely character at each time step. It is quick and it can create records that are near the first articulation.

**Beam search decoder:** A pillar search decoder with language model re-scoring permits checking numerous conceivable unraveling (shafts) without a moment's delay with relegating a higher score for more plausible N-grams as indicated by a given language model. The language model assists with rectifying incorrect spelling mistakes.

As shown in fig 3, Input files are downloaded from other sources that are related to our context, these files are then processed using NLTK library to extract noun.

These files are taken from internet which are then stored in a file for sentence segmentation. Sentence division is a procedure of deciding the more drawn out preparing units comprising of at least one word. This assignment includes distinguishing sentence limits between words in various sentences.

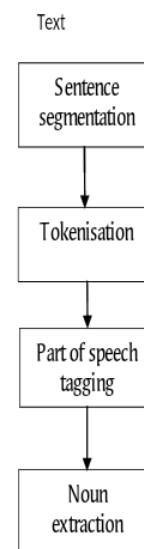
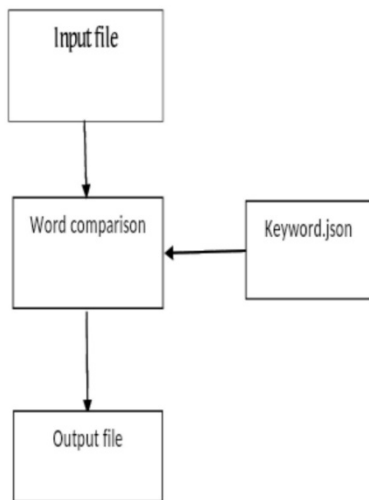


Fig 3: Extraction of noun

After sentence segmentation, it is split into smaller parts called as tokens. Data is secured here. Tokenization is the way toward separating the arrangement of characters in a book by finding the word limits, the focuses where single word closes and another starts. For computational etymology purposes, the words along these lines identified are as often as possible alluded to as tokens.

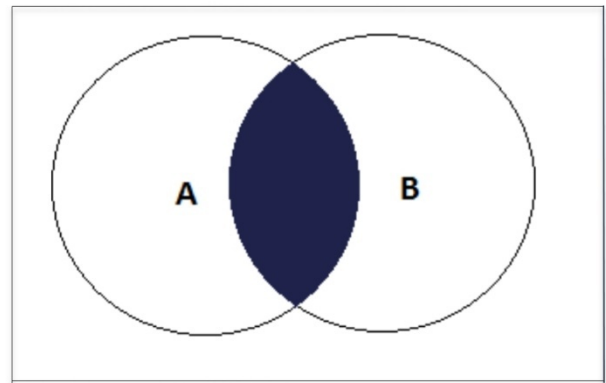
Grammatical form labeling is mind boggling task in substance extraction. The thought is to coordinate the token with the relating labels (things, action words, descriptors, intensifiers and so on.). The way toward characterizing words into their grammatical forms and marking them in like manner is known as a major aspect of discourse labeling.

Entity extraction refers to the procedure of recognizing and arranging key components from text into predefined classifications. It converts unstructured to structured data, which can be processed for retrieving information, extracting facts and question answering.



**Fig 4: Extraction of keyword**

As shown in fig 4, Input file is the text file extracted from the video by following above process. Now this file is taken as input for word comparison. Here words are compared to the json file. If the words are found to be matching then it is assumed to be a keyword and then stored in a output file.



**Fig 5: Intersection for extracting the keywords**

Where, A ->Output of downloaded files from internet for comparison.

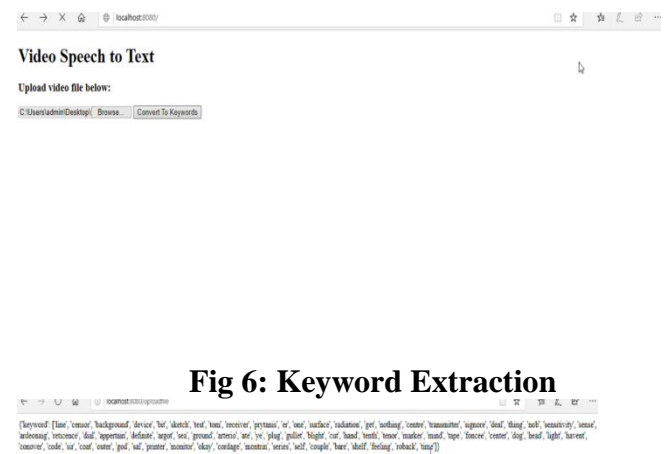
B ->Output obtained from final conversion (audio to text)

In the first part extraction of nouns from the text file downloaded from the source is done.

Later this is intersected with the document obtained from conversion of video to text. By the intersection of the set of nouns and the text file we get the required keywords.

**III. RESULT AND DISCUSSION**

This project is based on concept of virtual classroom. The aim of the virtual classroom is to create an environment that will help stimulate purposeful communication, which is to conduct class in absence of lecturer. Conducting class with the help of lecture video. In this project, extraction of keyphrases from the video is done.



**Fig 6: Keyword Extraction**

**Fig 7: Output**

The output for conversion of video to audio and then audio to text conversion is obtained. Firstly, the nouns extracted from the text file obtained from the source. The nouns are extracted by various steps like sentence segmentation, tokenization, part of speech tagging, and entity extraction using nltk.tree. After the noun extraction, extraction of keywords is done. This is done using word comparison and hence the second figure is obtained which are the set of keywords. The file is selected for the conversion, on clicking the convert button the output is obtained in keyword. Text file.

#### IV. CONCLUSION

In this paper, extracted keywords from video are done. However, presently only extraction of keywords is done. Since our idea of project was based on concept of virtual classroom in future, we will concentrate on getting the definition and explanations of the extracted words. The main objective of our project is to pick the main words from the video for easy understanding of the lecture. By using this method, lecture coming to class to conduct lecture can be reduced. By using such an automated system, the interference of manual labor can be reduced. Using this system, we can easily get set of keywords from the video.

#### V. FUTURE WORK

This project is based on concept of virtual classroom. The aim of the virtual classroom is to create an environment that will help stimulate purposeful communication, which is to conduct class in absence of lecturer. Conducting class with the help of lecture video. In this project, extraction of keyphrases from the video is done. In continuation to this project we can give the explanation of the extracted key words. Extracted keywords are stored in a file, now these keywords meanings are to be known. As the concept was about virtual classroom, when a lecture video is played it is necessary the student understand the lecture going on. When students have any questions on the related concepts, explanations should be in hand. Further elimination of noise in class can be done. Monitoring of class can be done to maintain discipline of the class in the absence of lecture.

#### VI. REFERENCE

- [1] "A Flexible Keyphrase Extraction Technique for Academic Literature", Gollam Rabby, Saiful Azad, Mufti

Mahmud, Kamal Z. Zamli, Mohammed Mostafizur Rahman Faculty of Computer Systems & Software Engineering, University Malaysia Pahang, Gambang, Kuantan, Malaysia. IBM Centre of Excellence, Gambang, Kuantan, Malaysia.

- [2] "Using Noun Phrase Heads to Extract Document Key phrases", Ken Barker and Nadia Cornacchia School of Information and Technology Engineering University of Ottawa, Canada
- [3] "Automatic Keyword Extraction for Text Summarization: A Survey", Santosh Kumar Bharti, Korra Sathya Babu, and Sanjay Kumar Jena
- [4] "Multi Feature audio segmentation for browsing and annotation" George Tzanetakis, Perry Cook, 35 Olden Street, Princeton NJ 08544, Olden Street, Princeton NJ 08544, USA
- [5] "Automatic Audio Sentiment Extraction Using Keyword Spotting", Lakshmi Kaushik, Abhijeet Sangwan, John H.L. Hansen Center for Robust Speech Systems (CRSS), Eric Jonsson School of Engineering, The University of Texas at Dallas (UTD), Richardson, Texas, U.S.A.
- [6] "Different Audio Feature Extraction using Segmentation", Gayatri M. Bhandari Research Scholar JJT University, Rajasthan.
- [7] "Automatic Keyphrase Extraction and Segmentation of Video Lectures", Arun Balagopalan, Lalitha Lakshmi. B, Vidhya Balasubramanian, Nithin Chandrasekharan, Ashwin Damodar, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India.
- [8] "Video Data Retrieval Based On Text Content Detection", Ms.M.Ayesha Parveen, Mr.P.Deepan2, Ms.S.Jummera Nasrin3, Ms.J.ISRATH BEGUM 4 Department of CSE, Arasu Engineering College, Kumbakonam, India.