Scientext: Enhancing Accessibility to Scientific Knowledge through Speech-to-Text Summarization

github: https://github.com/ayesha-mohsin/Scientext

Shaik Ayesha Mohsin || Areeba Hassan

Abstract:

This project aims to revolutionize access to scientific knowledge by developing a system capable of transcribing spoken scientific discussions from YouTube videos into text and generating comprehensive summaries. By integrating advanced speech recognition technology with state-of-the-art text summarization techniques, Scientext seeks to bridge the gap between scientific discourse and broader audiences. Through this initiative, we aim to democratize access to scientific knowledge and foster greater collaboration within the research community.

1. Motivation:

Scientific discourse serves as the foundation for innovation and progress within the research community. However, the accessibility of scientific discussions is often hindered by language barriers and technical complexities. By leveraging YouTube videos as a source of scientific discussions, Scientext aims to break down these barriers and make scientific knowledge more accessible to a wider audience. This initiative is driven by the recognition that facilitating access to scientific discourse is essential for driving innovation and addressing global challenges effectively.

2. Significance:

The significance of Scientext lies in its potential to democratize access to scientific knowledge and foster collaboration within the research community. By transcribing and summarizing scientific discussions from YouTube videos, Scientext empowers individuals from diverse backgrounds to engage with complex research topics effectively. This accessibility can accelerate the pace of discovery, drive interdisciplinary collaboration, and ultimately contribute to addressing societal challenges.

3. Objectives:

- Develop a robust speech recognition system capable of accurately transcribing scientific discussions from YouTube videos.
- Implement advanced text summarization techniques to distill transcribed discussions into concise and informative summaries.
- Evaluate the effectiveness of the proposed system through rigorous testing and user studies to ensure usability and accessibility.
- Provide a user-friendly interface for interacting with the system, allowing users to upload YouTube video links and access the generated transcripts and summaries seamlessly.

4. Features:

Scientext will offer the following features:

- Utilization of state-of-the-art speech recognition models, such as DeepSpeech or wav2vec, for accurate transcription of scientific discussions.
- Implementation of text summarization models leveraging techniques such as transformer-based architectures (e.g., BERT) to generate informative summaries.
- Integration of the speech recognition and summarization components into a cohesive pipeline, ensuring seamless processing of scientific discussions.
- Development of a user-friendly web-based interface allowing users to upload YouTube video links and access the generated transcripts and summaries effortlessly.

5. Dataset:

Scientext will utilize publicly available YouTube videos containing scientific discussions across various domains. Additionally, pre-existing datasets, such as TED Talks or academic lecture recordings, may be considered for training and evaluation purposes.

6. Workflow:

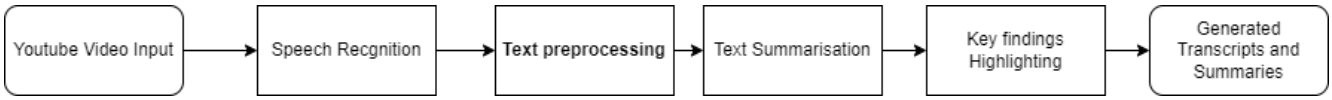

Fig1.Workflow Diagram

The workflow involves the following steps:

1. YouTube Video Input:
    - Users provide links to YouTube videos containing scientific discussions as input to the system.
2. Speech Recognition:
    - The speech recognition module transcribes the audio from YouTube videos into text, converting spoken words into written form.
3. Text Preprocessing:
    - The transcribed text undergoes preprocessing steps such as punctuation removal, tokenization, and stop-word removal to prepare it for summarization.
4. Text Summarization:
    - The preprocessed text is fed into the text summarization model, which generates a concise summary of the scientific discussion. Techniques such as extractive or abstractive summarization may be employed based on the project requirements.
5. Key Findings Highlighting:
    - The generated summary is enhanced by highlighting key findings, important concepts, and noteworthy insights to facilitate better understanding.
6. User Interface:
    - Users interact with the system through a user-friendly web-based interface, where they can upload YouTube video links and access the generated transcripts and summaries.

7. Conclusion:

Scientext represents a significant advancement in making scientific knowledge accessible to a broader audience. By leveraging YouTube videos as a source of scientific discussions and employing advanced NLP techniques, Scientext aims to democratize access to scientific discourse. Through this initiative, we hope to foster collaboration, drive innovation, and address global challenges effectively.