International Journal of Research Publication and Reviews

Automated Notes Maker from Audio Recordings

Manoj Kumar A[1], Janani P[2], Siva Subramanian G[3], Kumaragurubaran K[4], Sundari P[5]

[1]Student, Info Institute of Engineering, Computer Science and Engineering, NH-209, Kovilpalayam, Coimbatore, 641107
manojkumararumainathan@gmail.com
[2]Student, Info Institute of Engineering, Computer Science and Engineering, NH-209, Kovilpalayam, Coimbatore, 641107
jananiprasath314@gmail.com
[3]Student, Info Institute of Engineering, Computer Science and Engineering, NH-209, Kovilpalayam, Coimbatore, 641107
gsivasubramanian19@gmail.com
[4]Student, Info Institute of Engineering, Computer Science and Engineering, NH-209, Kovilpalayam, Coimbatore, 641107 *gurubaran677@gmail.com*
[5]Assistant Professor, Info Institute of Engineering, Computer Science and Engineering, NH-209, Kovilpalayam, Coimbatore, 641107
Sundarimuppal19@gmail.com

ABSTRACT:

Automated note-taking software uses machine learning and natural language processing to transcribe spoken words from audio recordings. In the current fast-paced environment, individuals and organisations regularly have to manage enormous amounts of data in the form of audio recordings, video recordings, and text documents. The manual summarization and transcription of such data can be laborious and time-consuming. This problem has been resolved by automated note-taking software for audio recordings. Natural language processing (NLP) and machine learning techniques are used by these technologies to transform spoken words into text, which is subsequently used to build a summary or set of notes. This project report looks at the effectiveness and utilisation of automated note-taking software for audio recordings.To transform the audio recordings of online classes into text-based PDF or Word documents, which are voice-based in correctness, a module for automated notes creator from audio recordings is shown in this presentation.

Keywords: Automated Notes Taking , Machine Learning, Artificial Intelligence, Notes Taker

1. Introduction

The essential component for advancement in the current sector is communication. Not just on a corporate level, but also on a personal level, it is crucial to convey information to the appropriate person and in the appropriate way. The globe and communication methods are both going towards digitization. In today's technologically advanced society, communication via phone calls, emails, text messages, and other channels has become essential. Many applications that function as a mediator and aid in efficiently transmitting messages in the form of text or audio signals over miles of networks have emerged in order to fulfil the objective of effective communication between two parties without obstacles. The majority of these apps make use of text to speech signal conversion. Come up with a method for converting voice-based recordings of online lessons into PDF or Word documents. Compare the spoken words correctness to the text in the PDF or Word document that was converted. Due to the voice-based nature of online classes, the initiative intends to deliver students text-based PDF/Word Documents of such classes.

2. Scope of the Project

Academic note-taking: Students can use automated note-taking software to record lectures, seminars, and other academic conversations. This can help them retain important knowledge so they can assess it later.

Business meetings: Professionals can utilize automated note-taking software to capture meetings and discussions. They can use this to record choices, deeds, and other important information.

Legal transcription: Law firms can employ automated note-taking software to transcribe court proceedings, depositions, and other legal talks. With this, they can produce accurate transcripts more quickly and efficiently.

Medical transcription: Healthcare professionals can use automated note-taking devices to record conversations about patients, medical conferences, and other issues related to healthcare. They may gain from this.

3. Requirement Analysis

3.1 System requirements

There are countless applications for automated note-taking from audio recordings. Here are several examples:

(To Build Automated Notes Maker)*

- Windows 8, 10 or Higher Versions

- RAM:4GB*

- Specification : I3 Processor or Higher (Similar to it)

3.2 Tools & Modules required

(Suggested)*

- Visual Studio Code

- Python (3.8+)

- Speech Recognition (3.9.0)

- Pydub (0.25.1)

- Python – docx (0.8.11)

- Matplotlib

- IO

- Streamlit (1.19.0)

- Github

4. Algorithm

Step 1: Start by importing the necessary libraries such as speech_recognition, pydub, matplotlib, python-docx, and Streamlit.

Step 2: Use Streamlit to create a user interface for the web application where users can upload their audio recordings.

Step 3: Allow users to upload their audio recordings through the web application. Perform a quality check on the audio file, ensuring it meets the required specifications (e.g., audio format, sample rate, etc.). Use the pydub library to handle audio file manipulation and checking.

Step 4: If the audio quality is satisfactory, segment the audio file into smaller parts (e.g., based on time intervals, silence detection, etc.) to process them separately. This step can improve transcription accuracy and efficiency.

Step 5: Apply the speech recognition module's recognizer to convert each audio segment into text format. Use the recognized library to handle speech recognition tasks and convert the audio segments into text.

Step 6: Combine the transcriptions from all the audio segments to form a complete transcript of the audio recording.

Step 7: Use the python-docx library to create a Word document (.docx) and write the transcribed text into it. Download the transcribed notes: Provide a download link or button on the web application, allowing users to download the generated Word document containing the transcribed notes.

Step 8: Optionally, you can use matplotlib to generate visualizations or summaries of the transcribed notes, such as word clouds, keyword extraction, or sentiment analysis.

Step 9: Start the Streamlit application, making it accessible to users through a web browser.

Step 10: Thoroughly test the application, including audio uploads, transcription accuracy, docx conversion, and downloading functionality. Debug your programme and fix any problems you find.

Step 11: Deploy the application to a web server or a cloud platform, making it publicly accessible for users.

5. Working manual of web app

Upload Audio File :We need to upload the audio file on web application to convert the audio file into transcribed notes document file.
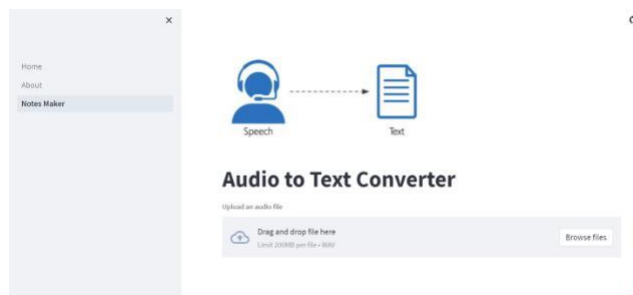
Converting to text : After uploading audio file , we just need to transcribe the audio to notes by using the option convert to text on web application , we can convert the audio file to text and it will take few seconds to convert (5-10 seconds approx)*
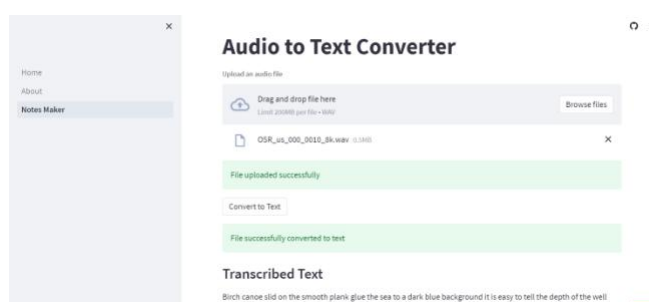
Download the notes : We can download the transcribed notes on docx format with the help of python-docx module. We can simply download it on web app after transcribing audio file.
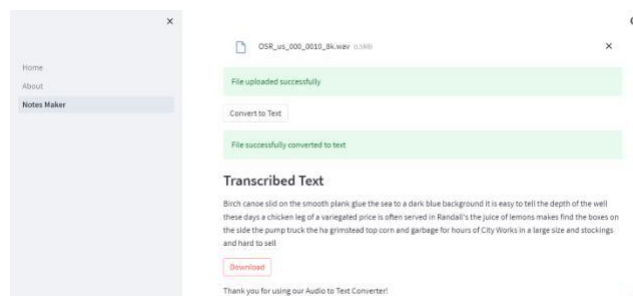
Limitations

- The machine learning models employed behind the scenes are not perfect, thus we cannot entirely rely on the translated or converted content.

- The current implementation performs a basic audio quality check, but it may not cover all potential issues. Enhance the quality check by implementing additional checks for factors like audio volume, background noise level, audio clipping, and audio artifacts to ensure better transcription results.

- Implement robust error handling mechanisms to handle failures gracefully, such as audio file upload failures, transcription errors, or network issues. Provide clear and informative feedback to users in case of errors or unsuccessful transcription attempts.

- The current web implementation only supports WAV audio files.

Future Enhancement

- Extend the project to support language translation, allowing users to transcribe audio recordings in one language and obtain the translated text in another language. This could involve integrating with machine translation APIs or implementing language translation models.

- To enhance the project's usability, you can consider adding support for other popular audio formats like MP3, FLAC, and OGG. This can be achieved by integrating additional libraries like ffmpeg or audioread to handle different audio formats.

Conclusion

Audio to text conversion systems are fantastic for assisting users in the examination and comparison of audio and video information. Additionally, these systems are able to produce text as notes in the appropriate forms. Future iterations of these technologies can increase their precision and personalisation.

S.No	Actions	Result
1.	Home Page	Yes
2.	Upload Audio	Yes
3.	Transcribing audio to text	Yes
4.	Transcribing in multiple languages	No
5.	Read text before download	Yes
6.	File download	Yes
7.	Accuracy (85%)	Yes

References

[1]   Benedict, Shajulin, Rahul Badami, and M. Bhagyalakshmi. "APM Bots: An Automated Presentation Maker for Tourists/Corporates Using NLP-Assisted Web Scraping Technique." In International Conference on Advanced Network Technologies and Intelligent Computing, pp. 647-659. Springer, Cham, 2021.

[2]   Speech to text conversion and summarization for effective understanding and documentation (https:// www.researchgate.net/publication/342147736_Speech_t o_text_conversion_and_summarization_for_effective_u nderstanding_and_ docum entation)

[3]   "BERT for Automatic Speech Recognition and Text-to-Speech Synthesis" (2020) by Jiatong Shi et al. This paper investigates the use of BERT, a popular language representation model, for speech-to-text and text-to-speech tasks.

[4]   "Audio-to-Semantic Vector Embeddings for Speech Summarization" (2021) by Amrith Setlur et al. This paper presents a novel method for audio summarization that uses audio-to-semantic vector embeddings to extract key information from audio recordings