IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12444

# AUTOMATED NOTE MAKER FROM AUDIO RECORDING

**Chaudhari Mahima[1], Mali Divya[2], Chaudhari Nehal[3], Kolhe Trupti[4], Ashish T. Bhole[5]**

UG Engineering student, Department of Computer Engineering, SSBT's College of Engineering and Technology, Jalgaon, India[1-4]

Associate professor, Department of Computer Engineering, SSBT's College of Engineering and Technology, Jalgaon, India[5]

**Abstract**: Today speech technologies are commonly available for a limited but interesting range of tasks. The technologies enable machines to respond correctly and reliably to human voices and provide useful and valuable services. As communicating with a computer is faster using voice rather than using a keyboard, people will prefer such a system. Communication among human beings is dominated by spoken language, therefore it is natural for people to expect voice interfaces with computers. can be accomplished by developing a voice recognition system - speech-to-text which allows the computer to translate voice requests and dictation into text. Voice recognition system - speech-to-text is the process of converting an acoustic signal which is captured using a microphone to a set of words. The recorded data can be used for document preparation. In the project is to able automated notes maker from audio recordings application based on AI, ML to make the computer understand speech commands and convert it into text and then PDF/WORD format.

**Keywords:** Automatic speech recognition, Technology, Automated captions, Qualitative data, Transcription**.**

## I. INTRODUCTION

In the present industry, communication is the key element to progress. Passing on information, to the right person, and in the right manner is very important, not just on a corporate level, but also on a personal level. The world is moving towards digitization, so are the means of communication. Most of the applications find the use of conversion from speech signals to text. The chapter related to automated notes maker from audio recording application. It involved building a system for Speech-to-Text using several classifiers available in the python library. Voice recognition is the task of capturing clear voice and convert it into text documented format. Motivation behind choosing the project because the variety of fields required Automated notes maker application in the modern life. In the system help in universities, companies, Organization etc. To construct the different applications, automated notes maker from audio recordings and techniques are needed. However, the techniques and methods are currently not available or only available in highly complex, expensive setups. The topic is to help solving the task in a simple setup. Such a solution would be of great importance and would be useful to the in general.

## II. BACKGROUND

In the last decade, automated captioning services have appeared in mainstream technology use. Until now, the focus of the services have been on the technical aspects, supporting pupils with special educational needs and supporting teaching and learning of second language students. Only limited explorations have been attempted regarding its use for research purposes: transcription of audio recordings. The system presents a proof-of-concept exploration utilizing three examples of automated transcription of audio recordings from different contexts; an interview, a public hearing and a classroom setting, and compares them against 'manual' transcription techniques in each case. It begins with an overview of literature on automated captioning and the use of voice recognition tools for the purposes of transcription.

## III. METHODOLOGY

### Algorithm/Steps

Primary database searching A primary sequence is one that has been experimentally determined, and a primary database is one that contain experimentally derived data. Searching a database in an efficient manner is a matter of prime importance. Methods that can run on small databases may not be effective with larger databases in terms of time and space efficiency.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12444

Following are the methods available in this Interface:
1. Start the application
2. Home Page
3. Selection of audio file
4. Upload audio file
5. Can play audio file
6. Convert text format
7. File will be downloadable into Pdf/Word format
8. File downladed
9. Stop

**Modules in Project**

Various Modules are Integrated into this application so as to perform various operations. Modules in the project are enlisted as follows

1. Uploading audio for converting into text - Audio file will be converted using deepgram library because of make an API request with a media file, and get a nice formatted transcript response
2. Converted text file into Pdf/Word - User can get the audio file into Pdf or word file format using the python pacakage pdfgen canvas.

## IV. RESULT & CONCLUSION

**Result:** So, the final result of the whole process as discussed above is as follows.
The project automated notes maker was successfully tested. Project converts the audio files into text format. As users choice it provides the choice of file format like Pdf/Word file.

| Sr.No. | Action | Result | |
|---|---|---|---|
| 1 | Home Page | Yes | - |
| 2 | Upload Audio | Yes | - |
| 3 | Upload multiple files | - | No |
| 4 | Transcription (English) | Yes | - |
| 5 | Transcription in different language | - | No |
| 6 | Play Audio | Yes | - |
| 7 | File downloadable in different format (e.g. pdf/word) | Yes | - |
| 8 | Fluency (clear/fluent) | Yes | - |
| 9 | Accuracy (80%) | Yes | - |

## CONCLUSION

The application unables user to take automated notes from audio recording the type of audio file required is mps. The deepgram open-source library is use in the application to convert audio recording into notes. The use of deepgram library to convert audio recordings into automated notes gives better accuracy of 80%. In future the work may be extended for other languages.

## REFERENCES

[1]. Bhupinder Singh, Neha Kapur, Puneet Kaur “Speech Recognition with Hidden Markov Model:A Review” InternationalJournal of Advanced Research in Computer and Software Engineering, Vol. 2, Issue 3, March 2012.
[2]. Hakan Erdogan, Ruhi Sarikaya, Yuqing Gao, “Using semantic analysis to improve speech recognition performance” Computer Speech and Language, ELSEVIER 200.
[3]. Chen, Jingdong, Yiteng Huang, Qi Li, and Kuldip K. Paliwal. ”Recognition of noisy speech using dynamic spectral subband centroids.” IEEE signal processing letters 11, no. 2 (2004): 258-261.
[4]. JiuHong, Wei, Wang LiPing, Li MengYang, and Wang YouWei. ”Advantages and deficiencies of the automated multimedia lecture recording system in lecture video production.” In 2009 International Forum on Computer Science-Technology and Applications, vol. 3, pp. 271-273. IEEE, 2009.
[5]. 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.