International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/IJSRCSEIT

Speech To Text Software is Called Live Chart

P. Ruhinaz Nawazi[1], Mrs. B. Jyothsna[2], Dr. R. Vasanthselvakumar[3]

M. Tech Student[1], Associate Professor[2], Professor[3]

Department of Computer Science and Engineering, VISWAM Engineering College, Madanapalle, Andhra Pradesh, India

ARTICLEINFO

Article History:

Accepted: 05 July 2023

Published: 21 July 2023

Publication Issue

Volume 9, Issue 4

July-August-2023

Page Number

172-177

ABSTRACT

In the grand symphony of human communication, an orchestra of technology has emerged, conducting a harmonious blend of convenience, efficiency, and inclusivity. At the forefront of this technological symphony lies speech-to-text live chat, a mesmerizing composition that orchestrates the seamless transformation of spoken language into written text. This abstract invites you to embark on a captivating journey into the realm of this revolutionary technology, as we explore its intricate composition, enchanting applications, and profound impact on society. Immersed within the digital realm, speech-to-text live chat draws upon the extraordinary capabilities of artificial intelligence, unraveling the secrets of speech recognition. Beneath its surface, the symphony conceals a symphony of algorithms and machine learning techniques, working tirelessly to transcribe spoken words into a tapestry of written text. Here, the magic of natural language processing intertwines with the intricacies of linguistic nuances, ensuring the faithful interpretation of our spoken melodies. Yet, this symphony's virtuosity extends far beyond the realm of technical prowess. Its transformative notes resonate across diverse domains, heralding a new era of accessibility and inclusivity. From individuals with hearing impairments yearning to unlock the harmonies of conversation to bustling conference halls echoing with real-time transcription, the symphony embraces all who seek to join its melodic cadence. Its ethereal timbre transcends barriers, enabling individuals to share their voices, unshackled by the constraints of auditory limitations. Nevertheless, every symphony encounters challenges on its path to perfection. We are using DTW(Dynamic Time Warping). This symphony of speech-to-text live chat confronts the intricacies of dialects, accents, and the relentless cacophony of background noise. Yet, amidst these challenges, it gracefully dances, striving to strike a delicate balance between accuracy and speed, between understanding and interpretation. With each performance, it refines its harmony, honing its ability to captivate audiences in awe-inspiring clarity. The societal impact of this symphony resounds with resplendent resonance. It orchestrates a transformation where the barriers that once separated individuals dissolve, replaced by a unified tapestry of shared understanding. In the virtuoso hands of this technology, the mute find their voices, the unheard find their platform, and the world amplifies its chorus of diversity. Through the symphony's triumphant crescendo, a new symphony of empathy and inclusion emerges, forever altering the fabric of human connection. As the curtains fall on this symphonic performance, the abstract casts a gaze towards the future, where the symphony continues to evolve. Enhanced accuracy, multilingual harmonies, and the integration with other communication tools beckon on the horizon. The symphony urges us to continue the symphony's composition, to refine its opus, and to unlock the full potential of its transformative melodies.

Keywords : Speech to text, Live chat

I. INTRODUCTION

In the vast realm of modern applications, Speech Recognition stands tall as a vital feature, enabling seamless interaction with technologies ranging from home automation to artificial intelligence. This article aims to be your guide, introducing the captivating world of Speech Recognition and delving into the capabilities of Python's pyttsx3 library. By harnessing the power of this library, we will explore the realm of real-time audio capture and recognition, immersing ourselves in a bi-directional stream that opens doors to a new level of communication.

II. RELATED WORKS

Huang Shan. Voice recognition systems in the telecom prepaid business applications [J]. Information Science, 2010.

Library plays an important role in student’s life. It is also helpful for the research scholars. Today libraries are established in all places. Finding a book in the library is a herculean task. Some people even loose interest while searching the required book. Searching for books in computer also takes much time. In our project we use voice recognition to find books which will make the herculean task easier. Even the people without system knowledge can access this and find books in an easier way. It reduces the time in searching books and makes the library user friendly.

Zhang Ping, Zhang Qiong. Based on HMM and BP neural network for speech recognition [J]. Cross-century, 2008.

This paper presents a study of automatic speech recognition system for Hindi utterances with regional Indian accents. In paper [3] we have designed matlab based ASR and control system for eight English key words by using simple rule base. This rule base algorithm is the beginning stage for Key Word recognition. In paper [4] we have designed Design of Hindi Key Word Recognition System for Home Automation System Using MFCC and DTW. Features of the speech signal are extracted in the form of MFCC coefficients and Dynamic Time Warping (DTW) has been used as features matching techniques. The recognition results are tested for clean and noisy test data. Average accuracy for clean data is 97.50 % while that for noisy data is 91.25 %. We face problem in noise environment to detect correct utterance now we are going to review different papers and find out different techniques to design our ASR control system for Hindi Key Words using MFCC and DTW in noise environment.

Yangshang Guo, Yang Jinlong. The speech recognition technology overview [J]. Computer, 2006.

We as humans speak and listen to each other in human-human interface, similar attempts have been made to develop vocally interactive computers, the computer that can give speech as output (speech synthesizer), given speech as (speech recognizer). Speech recognition allows the machine to turn the speech signal into text or commands through the process of identification and understanding, and also makes the function of natural voice communication. Developing and understanding Speech Recognition systems is an inter-disciplinary activity, taking expertise in linguistics, computer science, and electrical engineering and its ultimate goal is to achieve natural language communication between man and machine. This paper provides a survey on speech recognition and discusses the techniques that enable computers to accept speech as input and shows the major developments in the field of speech recognition. This paper highlights the speech recognition techniques and provides a brief description about the four stages in which the speech recognition techniques are classified

Yu Tiecheng. The current development of speech recognition [J]. Communication World, 2005.

This paper presents a brief survey on Automatic Speech Recognition and discusses the major themes and advances made in the past 60 years of research, so as to provide a technological perspective and an appreciation of the fundamental progress that has been accomplished in this important area of speech communication. After years of research and development the accuracy of automatic speech recognition remains one of the important research challenges (eg., variations of the context, speakers, and environment).The design of Speech Recognition system requires careful attentions to the following issues: Definition of various types of speech classes, speech representation, feature extraction techniques, speech classifiers, database and performance evaluation. The problems that are existing in ASR and the various techniques to solve these problems constructed by various research workers have been presented in a chronological order. Hence authors hope that this work shall be a contribution in the area of speech recognition. The objective of this review paper is to summarize and compare some of the well-known methods used in various stages of speech recognition system and identify research topic and applications which are at the forefront of this exciting and challenging field.

Ren Tianping. Application of speech recognition technology [J]. Henan Science and Technology, 2005.

This paper presents a study of automatic speech recognition system for Hindi utterances with regional Indian accents. In paper [3] we have designed matlab based ASR and control system for eight English key words by using simple rule base. This rule base algorithm is the beginning stage for Key Word recognition. In paper [4] we have designed Design of Hindi Key Word Recognition System for Home Automation System Using MFCC and DTW. Features of the speech signal are extracted in the form of MFCC coefficients and Dynamic Time Warping (DTW) has been used as features matching techniques. The recognition results are tested for clean and noisy test data. Average accuracy for clean data is 97.50 % while that for noisy data is 91.25 %. We face problem in noise environment to detect correct utterance now we are going to review different papers and find out different techniques to design our ASR control system for Hindi Key Words using MFCC and DTW in noise environment.

III. METHODS AND MATERIAL

Proposed system:

Despite the challenges encountered by speech recognition systems, there is immense room for improvement, making them more accessible and widely applicable. As voice recognition technology advances, it is foreseeable that speech recognition systems will become more sophisticated and their applications more extensive. A diverse array of speech recognition systems will emerge in the market, leading people to adapt their speech patterns to accommodate different recognition systems. In this context, we propose a speech-to-text recognizer utilizing Voice Manager, coupled with live chat functionality, to further enhance the capabilities and user experience of speech recognition technology.

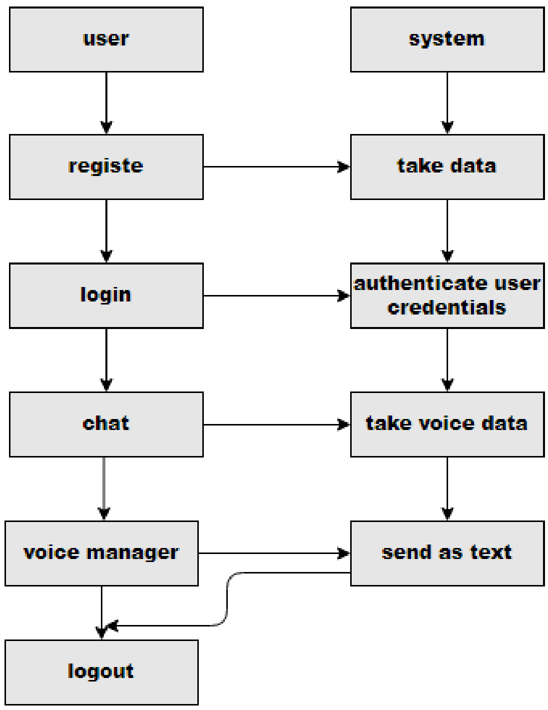


Figure 1 : Block diagram

IV. IMPLEMENTATION

Neural Network-based Models: Traditional speech recognition systems relied on statistical models, such as Hidden Markov Models (HMMs). However, with the advent of deep learning and neural networks, modern speech recognition systems employ deep neural networks, such as recurrent neural networks (RNNs) or convolutional neural networks (CNNs). These models have shown remarkable performance improvements by learning complex patterns and dependencies in speech data.

End-to-End Models: Traditional speech recognition systems consisted of multiple stages, including feature extraction, acoustic modeling, pronunciation modeling, and language modeling. However, end-to-end models directly convert spoken language into text without explicit intermediate steps. These models, such as the Listen, Attend and Spell (LAS) and Connectionist Temporal Classification (CTC) models, have simplified the speech recognition pipeline and achieved competitive accuracy.

Transfer Learning and Pre-training: Transfer learning techniques, such as pre-training on large amounts of labeled data, have proven effective in improving speech-to-text performance. Models like BERT (Bidirectional Encoder Representations from Transformers) and wav2vec leverage large-scale pre-training on vast amounts of unlabeled data to learn general speech representations, which can then be fine-tuned on specific speech recognition tasks. This approach has helped reduce the need for massive amounts of labeled training data.

Language Modeling: Language models, such as transformer-based models like GPT (Generative Pre-trained Transformer), have been widely used to enhance the quality of speech-to-text systems. By incorporating language models into the decoding process, the systems can utilize context and linguistic constraints to improve transcription accuracy, especially in cases of ambiguous or noisy input.

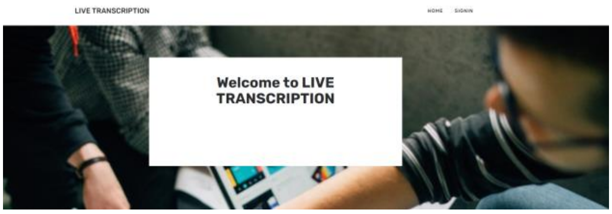
Data Augmentation and Domain Adaptation: To improve the robustness of speech recognition systems, techniques like data augmentation and domain adaptation are employed. Data augmentation involves artificially expanding the training dataset by applying transformations like noise injection, reverberation, and speed perturbation. Domain adaptation techniques aim to make the system more robust to variations in speakers, acoustic conditions, and languages by fine-tuning or adapting the model on target domain-specific data.

Real-time Processing: Advancements in hardware and software have enabled the deployment of speech-to-text systems capable of real-time processing. Through optimization techniques, such as model compression, quantization, and parallelization, models can be designed to run efficiently on devices with limited computational resources, including mobile phones and edge devices.

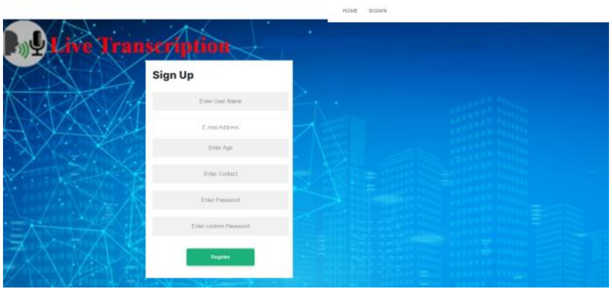
Multi-modal Integration: Modern speech-to-text systems can leverage multiple modalities, such as audio, video, and text, to improve accuracy and user experience. By incorporating visual cues from lip movements or sign language, these systems can enhance transcription accuracy, especially in noisy environments or for individuals with speech impairments.

V. RESULTS AND DISCUSSION

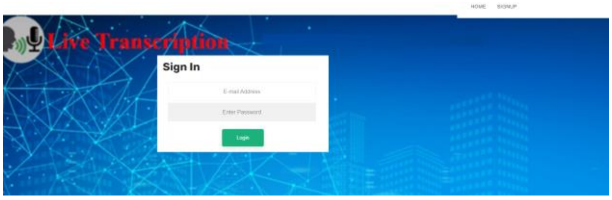
The following screenshots are depicted the flow and working process of project.



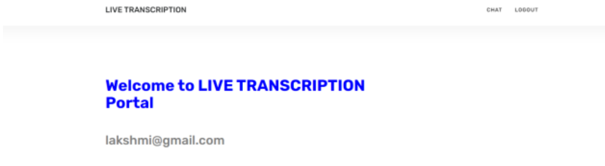
In above screen we have home page



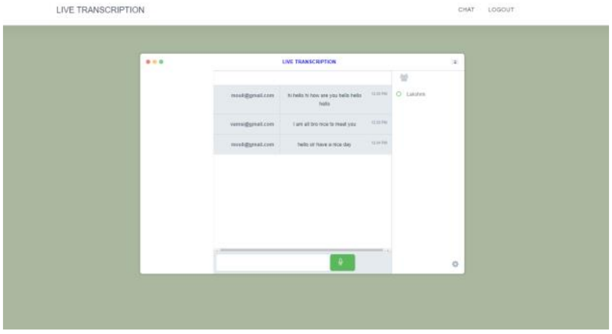
In above screen sign up page we signup with our details



In above screen we sign in with registered details



In above screen we have userhome page



In above screen we have chatting or live transcription page

VI. Conclusion:

The Speech-to-Text Live Chat project presents itself as a superior alternative to conventional communication methods due to its compelling advantages. Firstly, it facilitates real-time transcription, eliminating the need for manual note-taking or transcription services, resulting in significant time and resource savings. Moreover, it promotes accessibility and inclusivity by enabling individuals with hearing impairments to actively participate in conversations and express their ideas. Additionally, the advanced speech recognition technology ensures high accuracy and efficiency in capturing spoken language, minimizing the risk of misinterpretation or miscommunication. Furthermore, the project seamlessly integrates speech recognition technology with live chat functionality, enabling instant and interactive communication across multiple platforms. Ultimately, the Speech-to-Text Live Chat project revolutionizes communication by fostering efficiency, accessibility, and inclusivity in our daily interactions.

VII. REFERENCES

[1]. Yu Tiecheng. The current development of speech recognition [J]. Communication World, 2005.

[2]. Ren Tianping. Application of speech recognition technology [J]. Henan Science and Technology, 2005.

[3]. L A Liporace. Maximum Likelihood for Multivariate Observation of Markov Sources. IEEE. Trans. IT, 1982, 28(5): 729-734

[4]. Zhang Ping, Zhang Qiong. Based on HMM and BP neural network for speech recognition [J]. Cross-century, 2008.

[5]. Yin Peng, Li Tao, Wang Haibing. Intelligent neural network system composed of the principle in speech recognition. Mini-Micro Systems,2000,21(8):836-839.

[6]. Jiang Ming Hu, in the Yuan Baozong, Lin Biqin. Neural networks for speech recognition research and progress. Telecommunications Science,1997,13(7):1-6.

[7]. Huang Shan. Voice recognition systems in the telecom prepaid business applications [J]. Information Science, 2010.

[8]. Yangshang Guo, Yang Jinlong. The speech recognition technology overview [J]. Computer, 2006.

Cite this article as :

P. Ruhinaz Nawazi, Mrs. B. Jyothsna, Dr. R. Vasanthselvakumar, "Speech To Text Software is Called Live Chart", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 4, pp.172-177, July-August-2023.