https://doi.org/10.46335/IJIES.2022.7.8.15 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 84-88
International Journal of Innovations in Engineering and Science, www.ijies.net

Multilingual Speech and Text Recognition and Translation

Priyanka Padmane [1], Ayush Pakhale [2], Sagar Agrel [3], Ankita Patel [4], Sarvesh Pimparkar [5], Prajwal Bagde[6]

[1] *Professor, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
[2,3,4,5,6] *Student, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
.

ayushpakhale29@gmail.com

***Received on**: 11 June ,2022* ***Revised on**: 31 July ,2022,* ***Published on**: 03 August ,2022*

Abstract -*The automated translation of one human speech into another is referred to as "machine translation." The primary goal is to cross the linguistic gap between people from different cultures, neighborhoods, or countries. There are 18 major texts and ten scripts that are commonly used. Because the majority of Indians, especially remote dwellers, could really know, read, or write English, a better language translator is essential. Machine translation systems that convert one content into another will benefit Indians, allowing them to live in a more intelligent society without language barriers. Because English is a global language and Hindi is the native tongue spoken by the vast majority of Indians, we posit an English to Hindi autoencoder based on runs. (RNN), LSTM (Long-term Memory), but also attention processes "Machine translation" automatic object transcription of one basic language into another. The primary goal is to bridge a linguistic divide between different native languages, groups, or countries. There are 18 languages and ten commonly used scripts.*

Keywords - ***RNN, LSTM, Speech to text, text to Speech, Multi linguistic.***

I -INTRODUCTION

Language barriers are an issue in today's communication;therefore, we created this application to help. Speech recognition and text translation are mostly used to convert speech to text and text to speech in order to comprehend the language spoken by the user during conversation. As a result, a person can recognize the speech of another person. Language barriers are a problem in today's communication, so we created this app to help. Speech recognition and document translation are most commonly used to convert audio to text and automatic speech recognition in order to understand the vocabulary spoken by the customer during a conversation. As a direct consequence, a human could really recognize another person's speech. Machine translation is an ongoing since 1940. A google translate system converts text or speech from one human words to another. To convert a document or material from that other language out of our own speech, machine translation is required. It helps to break down linguistic barriers. NLP, or natural language processing, is an area of computer science that intends to bridge the gaps. The language understanding principle is simple, and it requires very little domain knowledge. To teach it to produce extremely long word sequences, a massive neural network was used. The model, in contrast to traditional machine translation, involves comprehensive phrase data sets and language models. The collaboration is in charge of the MT system's first condition antecedent. Translation is culturally important in ancient civilizations where different languages are spoken, so machine translation is significant. Furthermore, the notion of a long short - term memory is used. Hindi is India's most spoken and its primary international language, whereas English is spoken around the world and is therefore an internationally

https://doi.org/10.46335/IJIES.2022.7.8.19 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 1-6
International Journal of Innovations in Engineering and Science, www.ijies.net

recognized vocabulary. During British colonial period in India, English was adopted as a talking language. As a result, both English and Hindi have a lot of followers. As both a result, translating through one language to another presupposes use of a translator. In this section, we'll learn how to translate English to Hindi.

II -RELATED WORK

The research is centered on principle machine translation. It is based on a multilingual database but also corpus management solution. The system architecture's parser and geometrical tools validate the sentence construction of the language configuration before converting that to the target language. The technique outlined in the article [1] necessarily requires a thorough understanding of the grammatical structures of both the source language. Statistical machine translation makes use of statistics. This is based on the data theory concept. The translation is guided by the probability distribution. The technique proposed in this work [2] utilizes the Decision rule and numerical theory to reduce errors.

A hybrid technique combining principle and statistical machine learning is used for conversion. The architecture includes a coupler, parser, verb storylines tagger, phrase rules, reorder, syntactic database, and translator. In this project, a splitter divides the original text into words, while a parser examines the spelling as well as semantic structure. The declension tagger inflects nouns, adjectives, and pronouns to denote unique, plural, case, and gender. The source communication is then reorganized, and indeed the destination language would be translated utilizing lexical rules. The neurons Russian language is used in study [4]. This work discusses the architecture's coder, decoder, residue left connection, and other components.

DEEP LEARNING:

Deep Learning is a relatively new classification algorithm that has seen widespread adoption in a wide range of applications. It helps the system learn in the same way one which humans do, and train it to perform better. Deep Learning algorithms can represent features by integrating supervised and unsupervised learning. This ability is known as feature extraction.

A better machine translation system necessitates the use of various deep learning techniques and libraries. RNNs, LSTMs, as well as other algorithms are used to train the system that will transfer the sentence from source to target. Using the system is characterized and artificial neural networks is a good idea because it adjusts the system to enhance the translation system's accuracy when compared to others.

The advantages of Neural Machine Pronunciation (SMT) models require only a fraction of the memory required by these models.

- When series of linked corpora have become available, Fully Convolutional Nets outperform original state methodologies on shorter texts.
- In longer sentences, NMT approaches can be coupled with verb algorithms to resolve the rare-word problem.

I. PROBLEM STATEMENT AND OBJECTIVE

A. Problem Statement

The goal of our project is to automate the proposal so that it can achieve the language problem that exists both within and between countries. The aforementioned programmed will handle all aspects of the application. The goal of the proposed system is to invent a robot that could translate, convert text to speech, identify and respond, and obtain text. A tiny proportion of English words will be used to verify the hypotheses technique.

B. Objectives

- Our main goal is to bring together all of the many features, such as speech recognition, text translation, text synthesis, and text extraction from images, into a single, user-friendly programmed.
- voice production
- It's easy to use

II. COMPONENTS OF SPEECH TO SPEECH TRANSLATION

Speech recognition technology realizes the user's speech input and converts it to source language text. b) Translator, which converts text from one language to another, or technology which it translates recognized words. c) Speech synthesis, also renowned as text-to-speech biosynthetic pathways, is a technique that converts interpreted text into speech but rather synthesizes speech in the language of another person. In addition, the general engineering natural language, as well as user interface technology is critical in this monologue translation system.

https://doi.org/10.46335/IJIES.2022.7.8.19 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 1-6
International Journal of Innovations in Engineering and Science, www.ijies.net

III. APPLICATION

As their accuracy levels rise, translation solutions are increasingly being used in more parts of the industry, creating additional applications and enhance servo models.
Transcription for Business Applications in Industry Despite the fact that major companies including Google Translate and Software Translator provide near-real-time interpretations, a few really "domains" or areas require very precise data for training precise towards the domain in order to improve reliability and applicability. Because their computer models are trained on generic data, generic translators are of limited use in this situation. These applications are used by small, middle, and large businesses. Some companies offer multi-domain translation services, which are solutions that can be customized across multiple domains, whereas others offer translation solutions.

In the following fields, URL translation solutions are required:

- Technology but also software for government, defense, and military applications
- Health-care coverage
- Legitimate

IV. FLOW CHART

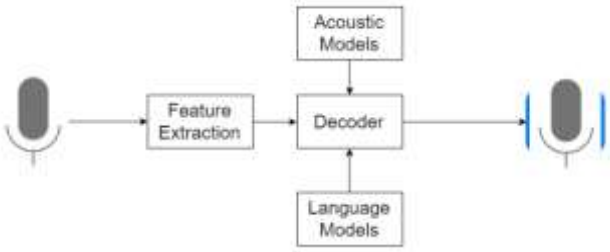

Fig 1 -Flow Chart

V. ARCHITECTURE

The user must first provide voice input during one of major languages: English, Kannada, Hindi, or Telugu, after which Look it up API translates speech to text. Then it proceeds to language translation, whereby the desired language is translated while the definition of the remark is preserved, which is treated by cognitive computing. After the translation software, it converts text to speech, and in the final stage, we get the interpreted voice output. The processes involved in speech-to-speech translation are as observes.

VI. FLOW CHART

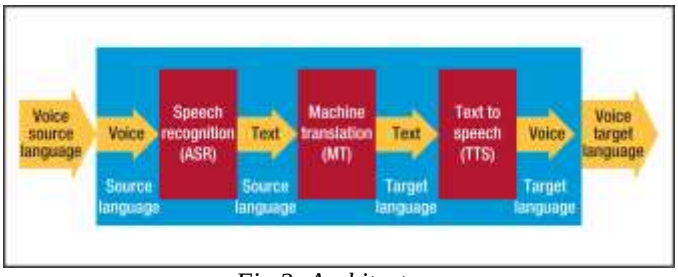

Fig 2- Architecture

VII. PROPOSED WORK

1)Input The voice input is taken as input to the system.

Voice input can be provided in any of four languages: Kannada, English, Hindi, or Telugu. Classification of Speech Recognition A variety of factors influence speech, including the speaker, throughput, and vocabulary size. Speech recognition is classified into several types based on length of both the speech, the presenter mode, and the size of the vocabulary. Classification of Speech Based on Utterances The four types of speech available are isolated word, connected text, ongoing speech, and spontaneous speech. Classification of Natural Language processing Based on Presenter Mode Approaches have been made are classified into two types based on speaker models: speaker relying and speaker independent.

2)Speech to Text Conversion Input from microphone:

Speech word samples from a sample will be extracted and separated into individual sentences. Using the microphone, the signal is recorded in the system. The input is human speech, which is sampled at a rate of 16,000 times per second. It should be capable of running in real time. Impairment generation systems, also known as tone output program needs, are enhanced and electronic mobile communication systems that assist people who have severe speech impairments in communicating verbally by having to replace their speech as well as writing. Speech generation devices are useful for people who struggle with verbal communication but since they allow someone to become active in conversations.
Spoken language dialogue systems include things like discourse synthesis and automatic speech recognition. Speech can be used to gather information as well. Although speech can be a more intuitive way of accessing information, controlling devices, and communicating, there may be other options: Speech may not be the most "natural" way to communicate with a computer. Fingers, eyes-free, faster, and

https://doi.org/10.46335/IJIES.2022.7.8.19 e-ISSN: 2456-3463
Vol. 7, No. 8, 2022, PP. 1-6
International Journal of Innovations in Engineering and Science, www.ijies.net

intuitive speech. The text output is produced following the voice processing.

3) Translation of Languages

Google's translation system primarily performs with text, but it also includes a built-in feature that allows you to pick up a speaker input but then play back the voice out over speakers. In the customary speech-to-speech translation approach, Google Translate uses a voice identifying number for text speech, text transformation, and voice recognition to create the audio linked with the text. Google's messaging service is a complicated candidate to surpass in terms of correct language translation due to it is very well deployment with massive amounts of information data from various sources. [13] A simple word replacement is used by a rudimentary mt system to convert information from one place to another.

4)Converting text to speech:

The text is the continuation of the recent phase. The original speech synthesizer will be used to convert text-to-speech. This correspondence will be easily understandable in the preferred language.

5) Voice output:

After the language translation, it converts text to speech and then actually creates voice output as the final stage.
The back end is just a search term that searches the following databases for data on front end:
The acoustic model is made up of acoustic noises that were trained to recognize different speech patterns.

Project Modules:

1. Login: -

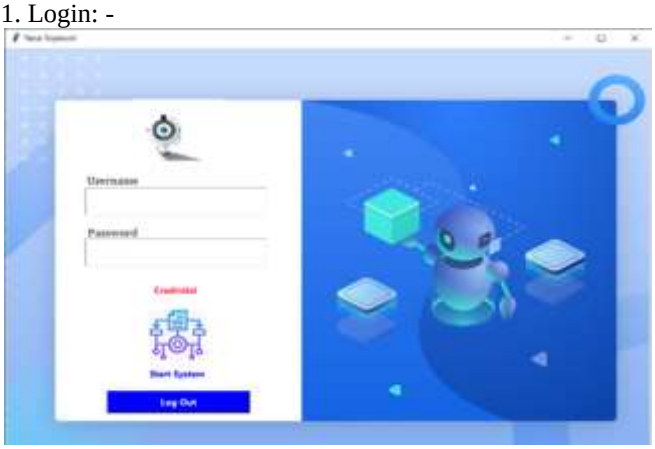

Fig 3 -Login Module

The user will log in to the system by entering his information, and our model will verify it using its credentials.

2. Home Screen: -

Fig 4- Home Screen

A user's home screen functions as a dashboard, with input options.

3. Audio and live mic input. :-

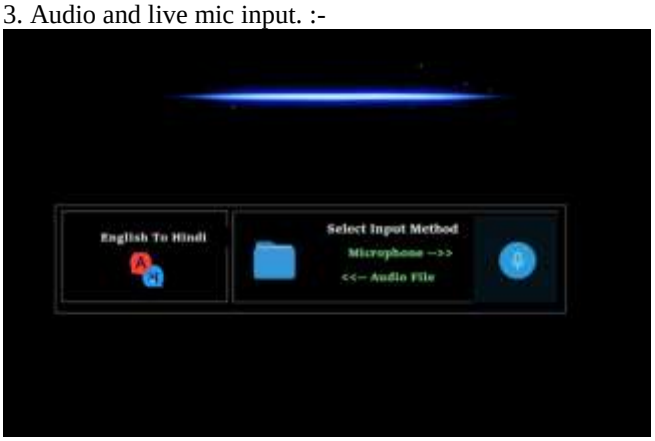

Fig 5- Audio and Live Mic Input

The input will be of two types: one will be pre-recorded audio, and the other will be live contribution in English.

A. Details of hardware and software

Hardware Requirements:

- Hard disk – 500 GB
- System – *I5* Processor
- RAM-4 GB

Software Requirements:

- LANGUAGE –
- Python
- Java

FRONT END: HTML, CSS

- APP- Java
- Web – python
- Database – SQLite
- Framework - flask

VIII. FUTURE WORK

This technology is continually being made for a desktop application, but it may one day be used on a cell device. As a result, rather than relying on a PC for language converting, users can make better use of this system by simply clicking a button on there own mobile device.

IX. CONCLUSION

We implemented the system for users with language barriers in this suggested system, and the interface is also user friendly so that users can easily engage with it. Because this application does not support the use of a dictionary to know the meaning of words, the user's task of understanding languages for communication is simplified.

REFERENCES

[1] *Manansala, S.K.Katti, "Speech Recognition by Machine: A Review", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 3, 2009 [2] Shyam Agrawal, Shweta Sinha, Pooja Singh, Jesper Olsen,"Development of text and speech database for Hindi and Indian English specific to mobile communication environment". [3] D.Sasirekha, E.Chandra, "Text to speech: a simple tutorial", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, And March 2012. [4] A. A. Tayade, Prof.R.V.Mante, Dr. P. N. Chatur,"Text Recognition and Translation Application for Smartphone.*

[2] *J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," arXiv, 2015.*

[3] *J. Redmon, A. Farhadi, "YOLO9000: Better, Faster, Stronger," arXiv, 2016.*

[4] *M. Swathi and K. V. Suresh, "Automatic Traffic Sign Detection and Recognition: A Review," 2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET), Chennai, 2017, pp. 1-6, https://ieeexplore.ieee.org/document/8186650 .*

[5] *S. Saini and V. Sahula. A survey of machine translation techniques andsystems for indian languages. In 2015 IEEE International Conferenceon Computational Intelligence Communication Technology, pages 676–681, Feb 2015.*

[6] *S. Chand. Empirical survey of machine translation tools. In 2016Second International Conference on Research in Computational Intelli-gence and Communication Networks (ICRCICN), pages 181–185, Sept2016.*

[7] *Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequencelearning with neural networks. CoRR, abs/1409.3215, 2014.*

[8] *Kyunghyun Cho, Bart Van Merri¨enboer, DzmitryBahdanau, and YoshuaBengio. On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259, 2014.*

[9] *International Conference onMachine Learning, pages 1310–1318,2013.*