Embedded systems and NLP Language Translation using Raspberry Pi

Vedansh Kumar, Bhavpreet Singh, Umrah Zargar, Arvind Kumar

SRM University, Delhi NCR Sonepat, Haryana, India

Corresponding author: Vedansh Kumar, Email: vedansh21.2000@gmail.com

Internet of Things, in general, is thought to be a dynamic and extensively allotted networked machine composed of a massive variety of identifiable clever gadgets. These gadgets can communicate and interact amongst themselves, and with end-users or different entities with inside the network. One of such a chip is called Raspberry Pi which is a fully customizable and programmable small computer board which can run complex tasks at a tremendous speed which is suitable for our tasks despite having a few disadvantages, the Raspberry Pi stays and less expensive pc with its very effective utilization in various variety of studies programs in IoT and Machine learning.One such use, considered here is its ability to send voice messages over the network through various methods. That in turn can be stacked up with Machine Learning and its various utility, one of the such being Natural Language processing. Understanding the words being conveyed to convert them into text stays it's one of the most used application.

Keywords: IoT (internet of things), NLP (natural language processing), Embedded systems.

2021. In Raju Pal & Praveen K. Shukla (eds.), *SCRS Conference Proceedings on Intelligent Systems*, 503–510. Computing & Intelligent Systems, SCRS, India.
https://doi.org/10.52458/978-93-91842-08-6-48

1 Introduction

The Internet of things can easily be considered the face of a new technological dawn. Nowadays all folks are using the web and therefore the use of small, cheap, and versatile element which allows end-user programming is additionally becoming popular and these small chips are perfect for the net of things. They need a broad horizon within the industry sector and might be used from large scale to small scale even in houses. Raspberry Pi, is one such chip with utility which spans as far as one's imagination allows, a fully customizable and programmable small computer board which can run complex tasks at an amazing speed which is suitable for our tasks despite having some disadvantages, the Raspberry Pi stays and fewer expensive pc with its very effective utilization in various style of studies programs in IoT and Machine learning. Raspberry PI was launched in 2012 and since its launch, it's been popular among all the folks who love to tinker with the tech. It can be operated the same way a customary PC could be used, requires a keyboard for entering commands, a display unit, and an influence supply. A credit card-sized computer with numerous features and a reasonable price of $25-35 is the ideal platform for connecting to a wide range of electronic devices. The vast majority of the system's components – its central and graphics processing units, audio and communications gear, and 256 MB of RAM (Fig 1. Model A) – 512 MB (Fig 1. Model B) microchip, are built onto one component. Raspberry pi shown in Fig. 2 is that the latest model which has models up to 8 GB ram and processor Quad-core cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz which might rival many small laptops and incorporates a powerful performance. The Raspberry Pi board is shown in Fig. 1, Fig. 2, and Fig. 3 contains essential (processor, the graphics chip, program memory - RAM) and other optional devices (various interfaces and connectors for peripherals). The processor of Raspberry Pi is a 32-bit, 700 MHz System on a Chip, which is made on the ARM11 architecture and may be overclocked and operating with certain methods for more power output. The unit is powered via the micro-USB connector while internet connectivity could also be via an Ethernet/LAN cable or via a USB dongle (Wi-Fi connectivity). [7]

Fig. 1. The RPI Model A and RPi Model B side by side [9]

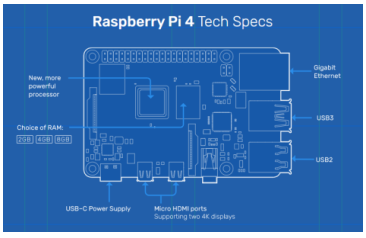

Fig. 2. Certain specifications of the latest model of Raspberry Pi [11]

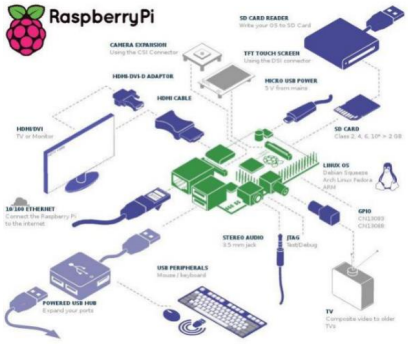

Fig. 3. Core-components of Raspberry Pi [10]

The Raspberry Pi, like any other computer that one may come across, utilizes an operating system to be operated. Raspberry Pi OS (formerly known as Raspbian), a Linux based option, is one of the best matches for Raspberry Pi because it's free and open-source, keeping the price of the platform low, and making it more available able for the general public who may not want to spend money on OS separately and learn it to operate on the other hand. Some non-Linux OS options are also available in the open market for the geeky ones to try out [8].

2 Natural Language Processing (NLP)

NLP is a field that combines computer science, artificial intelligence, and linguistics. The goal is that computers process the natural language or "understand" to perform activities such as translation. The symbols of the categorization of a language can be encoded in a variety of forms as a communication: sound, gesture, text, visuals, and so on. Any of these can be expressed in human language. Since human languages are ambiguous (unlike programming and other formal languages), there is a high degree of complexity in the representation, acquisition, and use of linguistic, situational, contextual, contextual, and visual knowledge of Human languages.

This branch has resulted in a rapid number of valuable applications. Vary in the complexity of complicated simple.

Here are some of them:

- Check spells, looking for keywords, and searches synonyms.
- Web data collection, such as product prices, appointments, places, people, and company names.
- Sorting: Read the shoulder text, the positive/negative mood in longer publications, etc.
- Automotive translation.
- Systems with which people are verbal.
- Answer difficult questions.

2.1 Machine Translation

The classic test of voice comprehension is a machine translation. Consists of two parts: linguistic analysis and linguistic creation.

We need to use a parallel corpus in a standard machine translation system - a collection of texts, each translates into one or more languages as the original. For example, given the source language (for example, French) and the target language e (for example, English), we must construct multiple statistical models, including a probabilistic formulation based on the Bayesian rule, a translation model 'p' trained on the parallel corpus, and a language model p(e) trained on the English-only corpus. Needless to say, this approach overlooks a slew of essential aspects, necessitates a great deal of human feature engineering, entails a slew of various and independent machine learning challenges, and is a complex system overall.

2.2 Types of Machine Translation in NLP

- SMT - Statistical Machine Translation
- RBMT - Rule-based Machine Translation
- HMT - Hybrid Machine Translation
- NMT -Neural Machine Translation

The central idea behind the idea of the end-to-end neural networks is the use of neural networks to provide a stand-alone translation of natural languages. To convert sequences into sequences, the neural machine translation often uses an encoder-decoder panel. The sequence-to-sequence models

(SEQ2SEQ) can be used for a variety of NLP activities, for example. B. Summary of the text, voice recognition, modeling of the DNA sequence, included.

2.3 Sequence to Sequence(seq2seq) Model

Fig 4. Fundamental principle of Sequence-to-Sequence Modelling [11]

As we aim to translate certain sentences from one language to another. Entry and exit are an active part of the sentences. In other words, sentences are a series of words that come in without following a pattern and still make sense at the end. The sequence-to-sequence model is based on this concept. Fig. 3 shows the process.

A standard Sequence-to-Sequence model has 2 major components:-

a) Encoder
b) Decoder

Both these parts are essentially two different recurrent neural networks (RNN) models combined into one giant network:

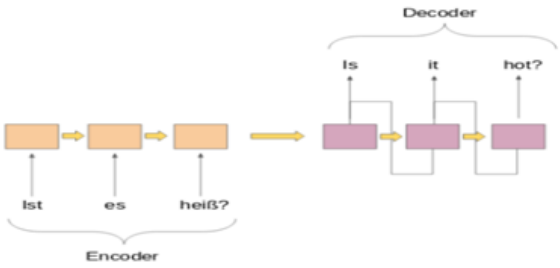

Fig. 5. Example of German-English encoder-decoder [11]

For example, the encoder-decoder structure offers a vector representation of each German word and then iteratively going from left to right through a recurrent neural network to generate a vector representation of the entire English sentence.

We call the recursive neural network used by the source language encoder, that is, to encode the sentence of the source language in a continuous and dense vector of real numbers.

Then a second recursive neural network is used on the target language to recursively decode the vector of the source language phrase, resulting in the desired English sentence. The whole decoding process takes place word by word and the decoding process ends when the ending part of the mentioned sentence is generated. The decoder is basically the RNN used by the target language. Note that each and every newly created English word is used as former supplied data to generate the next English term. Therefore, the decoder can be seen as a language model containing the target language information of the source language.

The encoder used in attention mechanism-based neural machine translation is quite different, with the aim of generating a vector representation containing aggregate information about each word of the source language, rather than a vector representation for the whole source language.

The advantage this approach provides us with is that the vector representation of each word in the source language contains the former and the latter contextual information. The decoder automatically finds contextual meaning from the source language associated with each target language word on all six sides of the target language.[1][2][5]

For example, when we are trying to translate a sentence which is in Hinglish, “Kya ye garam hai?” basically should translate to “Is it hot?” in English. Also, the Hindi sentence “□□□□ □□ □□□□ □□” should translate to “Is it hot”. But if the machine starts translating it word by word, then it will turn out to be “What this hot is?” literally in English. The approach here is to see if multiple words turn out to be actually one in the target language, “Kya ye” becomes “is it” in English and “garam hai?” is converted to “hot?” because the machine understands that “hai” means the end of the sentence and needs not to be included after translation as such.

As a result, the attention mechanism changes the way information is conveyed and can dynamically determine the most appropriate context, allowing it to better address the challenges of long-distance communication and significantly increase the efficiency of the resulting neural machine translation. As a result, the attention-based encoder-decoder model has become the most widely used means of neural machine translation.

3 Techniques

Speech translation is the process of converting/translating a spoken sentence from a language to another. The purpose of speech conversion/translation is to improve the man-to-man conversation. It should provide fast and accurate translations. An embedded inexpensive gadget is being used to carry out language translation like the Raspberry Pi.

There are three stages to the language translation process.

- The first step is to transcribe the input voice-over.
- The next step identifies the computer's language and deals with the audio file.
- The last step is to convert audio to text using online translation. The audio is being uploaded to an internet transcription service, and the language must be translated via a web server.

The RPI Model B is being used in this system. It's employed in the speaker verification procedure. It is utilized by RPI to recognize and translate the language, and it must be optimized. It's for RPI, and it's used to determine the language of the speaker. The language which has been detected is returned to the Raspberry Pi and used for speech/ voice recognition. Google Speech API is utilized to power the speech/voice recognition as well as translate the target language.

In the end, the player must generate a result. Only five languages have been translated into this system. Hindi, Spanish, English, French, and Chinese are the languages spoken. For this, Google's Speech to Text service was the best choice for RPi's speech recognition needs.

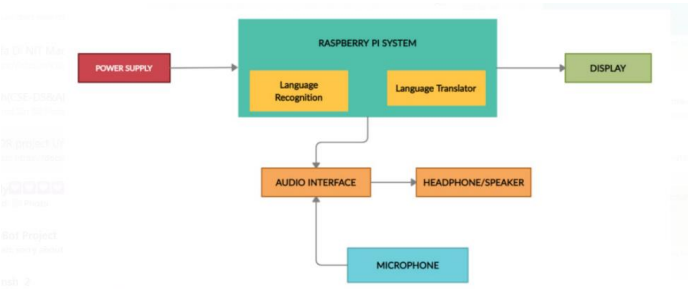

Fig. 6. Use Case diagram for our model

The Raspberry Pi runs this code to send a .wav file to the specified URL. An API key may be obtained from Google's Developers Console to utilize the voice recognition service; this API key is required to use Google's speech recognition service.

After that, the text which is in the source language is converted/translated to text which is in the target language using the Microsoft Translator. The audio file created from the text conversion is placed in the player's mp3 library. Using an mp3 player, you may listen to the finished product. When it comes to language translation, 15 different languages are needed to accomplish this. These are the languages: English, Hindi, Spanish, Chinese, Tamil, Danish, French, German, Italian, Japanese, Malay, Arabic, Turkish, Polish, and Urdu.

4 Applications of machine translation using embedded systems in the project

4.1 Text translation

In a variety of sentence and text-level applications of translation, automated text translation is widely employed. The translation of recovery, as well as inquiry inputs and Optical Character Recognition translation, are all examples of sentence-level translation applications.

4.2 Speech translation

The use of voice input has grown in popularity as a means of human-machine communication, and the translation of speech has emerged as a crucial use case, thanks to the rapid expansion of mobile applications. "Source language discourse source language text-target language text-target language discourse" is the basic cycle of discourse interpretation.

In this cycle, the translation of coded text from one language to another is crucial. Furthermore, ASR and text-to-speech (TTs) are required on both the front end and back end.

5 The Model

Using a Raspberry pi model 4 with 8 GB ram we can start the process of real-time language translation. The power supply used is 5v for this model and we're also using peripherals i.e., Audio input device (Microphone), Audio output device (Bluetooth Receiver for wireless audio transmission). The incoming audio waves is converted into electrical signals and then the computer process the signal by a method known as sampling. After sampling, we'll process the incoming sound into a waveform and then feed this waveform into an algorithm running on raspberry pi. Then using a simple code, the audio wave file is sent to Google Speech API which can convert the speech wave to text and then we can use another API, IBM Watson to translate the given text and then again convert it into speech using the same IBM Watson and finally the audio is returned to raspberry pi hence to the end-user. As IBM Watson, and Google Speech API supports many languages (119 languages) we can easily translate new languages using this technique. Also, there will be a passive listening feature that always looks for a specific keyword said by the user i.e., "Translate This" and then we can start the whole recording and conversion process.

6 Conclusion

Just by using a very small in size board with a few wires and the right amount of work done on code, wonders can be created. With this little combination of machine learning with this new and upcoming circuit name Raspberry pi, things like the language barrier can be overcome, which is one of the things that stops many people to explore the beautiful world we live in. By implementing this type of model using these techniques we will be able to translate languages even in real-time if the specific efficiency level is met. Multiple languages, that span from internationally accepted languages that the majority of the population is familiar with to the regional languages that give even the locals a hard time communicating due to various reasons, can be used for translation purposes and this model will also be portable as to suffice the modern need of wireless and small technology.

References

[1] Zong, Z. and Hong, C. On Application of Natural Language Processing in Machine Translation. In the *3rd International Conference on Mechanical, Control and Computer Engineering*, 506-510.
[2] Le et al. (2016). Neural Network for Machine Translation, at Production. https://research.googleblog.com/2016/09/a-neural-network-for-machine.htm.
[3] Bahdanau, D., Chok and Bengio, Y. (2014). Neural Machine Translation by jointly learning to align and translate. *arXiv*:1409.0473.
[4] Sharmila, R. and Malliga, R. (2017). Language Translator Using Raspberry Pi. *Asian Journal of Applied Science and Technology,* 1(3):281-282.
[5] Jurafsky, D. and Martin, J. Speech and Language Processing. *Pearson Education Limited, 2nd Ed, 944 pp. ISBN-10: 1292025433.*
[6] Maksimović, M. et al. (2014). Raspberry Pi as Internet of Things hardware: Performances and Constraints. In the *Conference: IcETRAN.*
[7] https://www.raspberrypi.org/products/raspberry-pi-3-model-b/
[8] https://www.raspberrypi.org/documentation/computers/raspberry-pi.html
[9] https://www.oreilly.com/library/view/raspberry-pi-cookbook/9781449365288/ch01.html
[10] https://fossbytes.com/getting-started-with-raspberry-pi-get-everything-youll-need-part-2/
[11] https://www.raspberrypi.org/products/raspberry-pi-4-model-b/specifications/