Machine Translation

Lyla B. Das*, Member, IEEE, Raghu C V**, Member, IEEE, Jagadanand G***, Member, IEEE,
V Sai Kiran', N Naveen Chandra Teja", V V V S N Murthy"', V Haswin Sai""
Department of Electronics and Communication Engineering
National Institute of Technology Calicut, Kerala - 673601, India
Email: *lbd@nitc.ac.in, **raghucv@nitc.ac.in, ***jagadanand@nitc.ac.in, 'vaikuntamsaikiran91@gmail.com,
"naveencteja@gmail.com, "'vallurivvsnmurthy@gmail.com, ""haswinvalluri@gmail.com

***Abstract*—Natural Language Processing (NLP) is a part of Artificial Intelligence (AI) which helps the machines to understand the human language and also it helps to manipulate that language. Within this NLP, there are a wide range of applications. One among them is Machine Translation(MT), which is a very important task in NLP. Machine Translation is a task of converting source language to target language automatically. This paper describes the MT task from English Language to Telugu Language. This can be used by people who don't know Telugu language, but they want to communicate in telugu. We have done this task using three different models. All these three models are Encoder Decoder Models. Firstly, the encoder takes the source language text as its input, learns the dependencies between the words and forms a representation vector. Then the decoder takes this vector and outputs target language words. The details about these models are presented Abstract—Natural Language Processing (NLP) is a part of Artificial Intelligence (AI) which helps the machines to understand and process human language for various purposes. The wide variety of applications for NLP is truly remarkable.**

One application is Machine Translation(MT), which is a very important and useful task. Machine Translation is the task of converting a source language to target language automatically. This paper describes the MT task from English to Telugu . This can be used by people who don't know Telugu , but want to communicate in Telugu. We have done this task using three different models. All the three models are of Encoder Decoder architecture ,but with progressively newer and better aspects . The encoder takes the source language text as its input, learns the dependencies between the words and forms a representation vector. Then the decoder takes this vector and outputs the words of the target language . The details about these models are presented. . We have also used a Bilingual Evaluation Under Study (BLEU) score to evaluate the performance of the models.

I. Introduction

As human beings we can write, speak and read many languages. But computers can understand only machine language which is in the form of zeros and ones .Natural Language Processing (NLP)is the technique to make computers understand the natural language used by us humans. NLP is now a very important subset of AI which uses many algorithms and tools of Computer Science. . NLP techniques are used to analyse the data that is created online and stored in databases. NLP has a wide variety of applications like sentimental analysis, text simplification ,text classification etc. Our work ' Machine Translation' is a very important application of NLP. Machine Translation is a task to do automatic translation between different languages. Generally ,translation is one of the most difficult tasks in the NLP area, because effective translation has to understand the exact meaning and tone of the input language and translate it to the target language with the same meaning and desired impact. Machine Translation can be done in three ways: Rule based MT, Statistical MT and Neural MT. In our work , we have Neural Network based Machine Translation.

II. Literature Survey

The project started by surveying the previous research made on Machine Translation. Based on this we are able to find that neural machine translation achieves significant improvement over primitive techniques: Rule based and Statistical Machine Translation. Ilya et al., 2014 [1] proposed an encoder-decoder model both implemented using LSTMs. The encoder reads the input sequence, each word at once to produce a context vector which is the representation of words seen so far and the decoder takes this vector and produces the target sentence. Gaurav Tiwari et al.,2020 explains clearly about Neural machine translation using Long short term memory networks along with attention (LSTM) [2]. The Encoder - Decoder model uses Bahdanau attention also known as additive attention. Bahdanau et al., 2016 [3] proposed the use of Bahdanau attention mechanism in encoder decoder architecture to mainly focus on important words seen so far while predicting the next word of the target sequence. Ashish Vaswani et al., 2017 proposed the transformer model in the paper attention is all you need [4]. Again this is also an encoder decoder model. These transformers consist of stacked layers of encoder and decoders. Ryan David Cunningham et al,.2021 clearly explained the fine tuning of the MarianMT model on the data [5].

III. Methodology

A. Data Collection

We need a dataset that consists of English sentences and their corresponding Telugu translations. The English to Telugu translation dataset we used in our project for training is taken from Samanantar indicNLP website. Samanantar is the largest publicly available parallel corpora collection for Indian languages. Coming to the test dataset we have taken 5k sentences (English and their corresponding Telugu text) ,chosen randomly from the internet.

B. Data Preprocessing

First of all, from the training data we have checked the translated sentences and we have removed the sentences which are not correct translations. After that we have used the NLTK library for preprocessing the data that includes removing html tags, question marks, exclamation marks and also converting the data to lowercase. For each sentence we have added certain tokens (start,end,pad). Data Preprocessing First of all, from the training data we checked the translated sentences and removed the sentences which are not correct translations. After that we used the NLTK library for preprocessing the data that includes removing html tags, question marks, exclamation marks and also converting the data to lowercase. For each sentence we have added certain tokens (start,end,pad).

C. Model

We have used the Encoder - Decoder architecture. The Encoder reads the input and summarizes the information in the form of a 'context vector' and the decoder takes this vector and outputs the target text. We did this translation using

1) Encoder - Decoder without attention - Bidirectional LSTM
2) Encoder - Decoder with attention - Bidirectional LSTM
3) Transformer based Encoder - Decoder model
4) Fine tuning of Pretrained MarianMT mode

1) Encoder - Decoder without attention - Bidirectional LSTM: .

The practice of permitting any neural network to retain sequence information in both backwards (future to past) and forwards (present to future) orientations is known as bidirectional long-short term memory (past to future). Our input runs in two directions in a bidirectional LSTM, which distinguishes it from a conventional LSTM. Using a standard LSTM data flows only in one direction, either backwards or forwards. However, with bi-directional LSTMs, we can have the information flow in both directions, preserving both the future and the past.. In the encoder part, the input text is encoded into a single context vector. Then the decoder takes this context vector and predicts the target language text. We used the Teacher Forcing concept in this model. Teacher forcing is the use of actual or expected output from the training dataset at the current time step y(t) as input in the next time step X(t+1) , rather than the output generated by the network. That is instead of giving output of the previous LSTM cell as current cell input, the ground truth value from the training set is applied. Fig 1 illustrates this model.

2) Encoder - Decoder with attention - Bidirectional LSTM: .

In this model we used an extra attention layer. Attention is a mechanism that was used to improve the Encoder - Decoder Bidirectional LSTM's machine translation performance. The Encoder-Decoder model's limitation of encoding the input sequence to 'one fixed length vecto' from which to decode each output time step is addressed with 'attention'. The attention

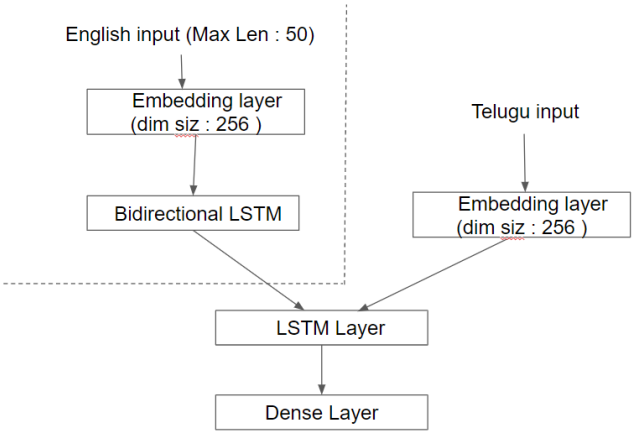

Fig. 1. Model using bidirectional LSTM

model creates a context vector that is filtered individually for each output time step, rather than storing the input sequence into a single fixed context vector. Using a set of attention weights, the decoder will be informed at each decoding step how much "attention" has to be paid to each input word. These attention weights supply the decoder with contextual information for translation. We used the 'Bahdanau attention layer' in our model in the decoder part. Figure 2 shows the architecture of this model.

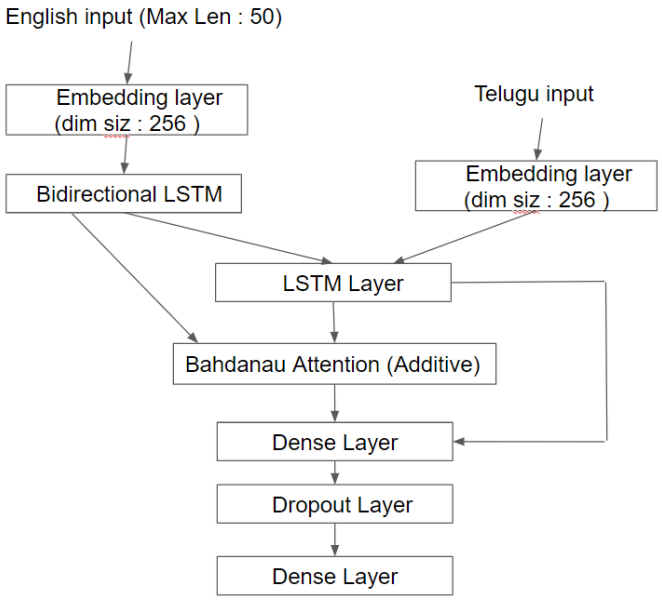

Fig. 2. Model using bidirectional LSTM with attention

3) Transformer based Encoder - Decoder model: .

It's made up of layers of encoders and decoders stacked on top of each other. The Encoder and Decoder stacks each have their own Embedding layers for each of their distinct inputs. All of the encoders are exactly the same. Likewise, all of the Decoders are the same. We have added positional encoding using sine and cosine functions. This model contains three

encoder and three decoder layers and eight self attention heads.Architecture of this model is shown in fig3 and fig4. 3) Transformer based Encoder - Decoder model: . IT is made up of layers of encoders and decoders stacked on top of each other. The Encoder and Decoder stacks each have their own Embedding layers for each of their distinct inputs. All of the encoders are exactly the same. Likewise, all of the decoders are the same. We have added positional encoding using sine and cosine functions. This model contains three encoder and three decoder layers and eight self attention heads.The architecture of this model is shown in Figures 3 and 4.

4) Fine tuning of Pretrained MarianMT model: .
Currently Transfer Learning is becoming Very popular. Transfer learning is the concept of reusing a pretrained model - we can fine tune this model on a small dataset also. The Marian-MT model we used in our project has an embedding size of 512, with 8 attention heads and six encoder and six decoder layers. This model is already trained on the OPUS dataset. Now we have fine tuned on our Samanantar dataset. The architecture of this model is similar to the transformer model shown above except that it has more layers.

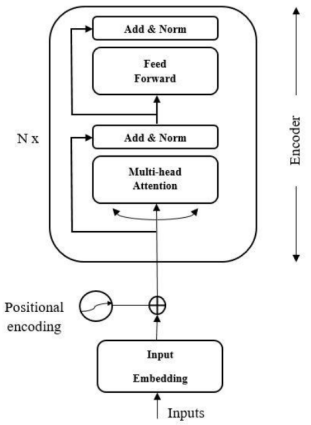

Fig. 3. Encoder of Transformer Model

IV. Addition of Voice Translation

This is an addition to normal translation and voice communication between persons is much welcome. There are people cannot read and write in a particular language. ,Such people can can use this for communication. Voice Translation is the process of translating conversational spoken word(s) and/or phrase(s), with the outcome shown on a screen or spoken aloud in the second language. The key to making this technology more efficient is to automate the procedure and make it an inter-audio conversion, which delivers simultaneous results without the need to start the translation process physically. People can just put the device on their heads and hear native speech in their own languages.

In our work, first converted speech to text and then translated it to target language and finally converted it into speech.

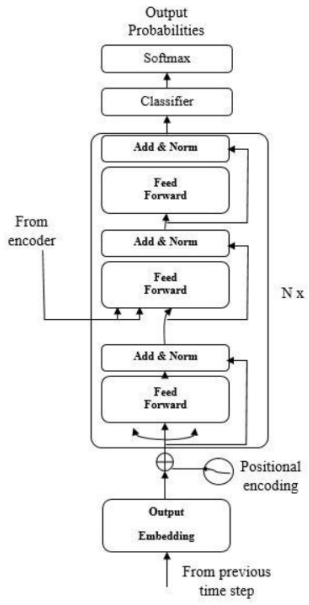

Fig. 4. Decoder of Transformer Model

For recognizing the words in speech we used Facebook's speech recognition model. This text is translated using a previously developed model to the target language. And again for converting back to speech, we used Google text to speech APIs.

V. Translation Bot

A bot is a software programme that automates and repeats pre-defined operations. Bots are designed to resemble or replace human user activity. Since they are automated, they can do the same work in less time than humans. Bots can be employed in sectors such as customer support, business, scheduling, search functionality, and entertainment. Bots normally communicate across a network. They connect with one another through internet-based platforms including instant messaging (IM), Twitterbots, and Discord bots. In this project we have done a discord bot. If we send an english message to this bot then it will translate it into telugu and send that telugu message back to us. We have inserted our developed model in this bot. This model translates from english to telugu language. As of now this bot only translates text messages. We can further develop this bot to translate voice messages also.

VI. Results And Conclusion

We trained each of our models for 100 epochs with Adam as its optimizer and learning rate of 1e-5. We also calculated BLEU scores for all the four models. This is tabulated in Figure 5. We find the best score to be obtained with the pre-trained model . Figure 6 shows a sample of the correct text translation done by our best model and Fig 7 shows the result of the translation done by the bot. We conclude by listing our

work as follows: A real time machine translation model from English to Telugu has been developed. This model is effective in translating English text to Telugu . We have also used this model to translate voice input. A translation bot has also been developed that is user friendly. Anyone can easily use this bot to communicate quickly. We are continuing this work to get the bot to translate voice messages as well.

Model	BLEU Score
Encoder - Decoder without attention - Bidirectional LSTM	0.685
Encoder - Decoder with attention - Bidirectional LSTM	0.743
Transformer based Encoder - Decoder model	0.83
Fine tuning of Pretrained MarianMT model	0.89

Fig. 5. BLEU scores for models

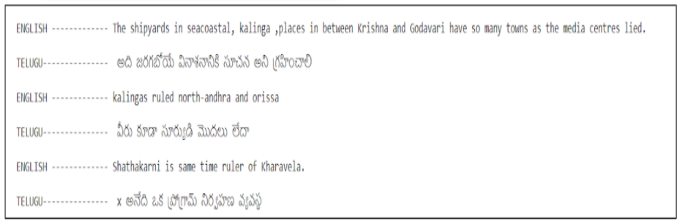

Fig. 6. Incorrect Translations

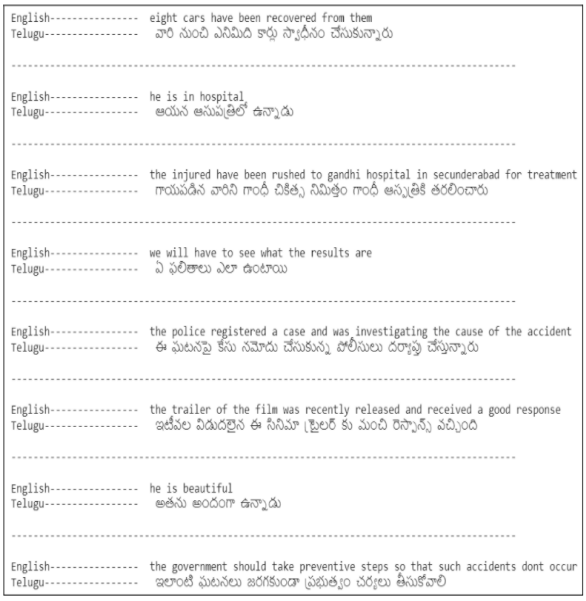

Fig. 7. Correct Translations

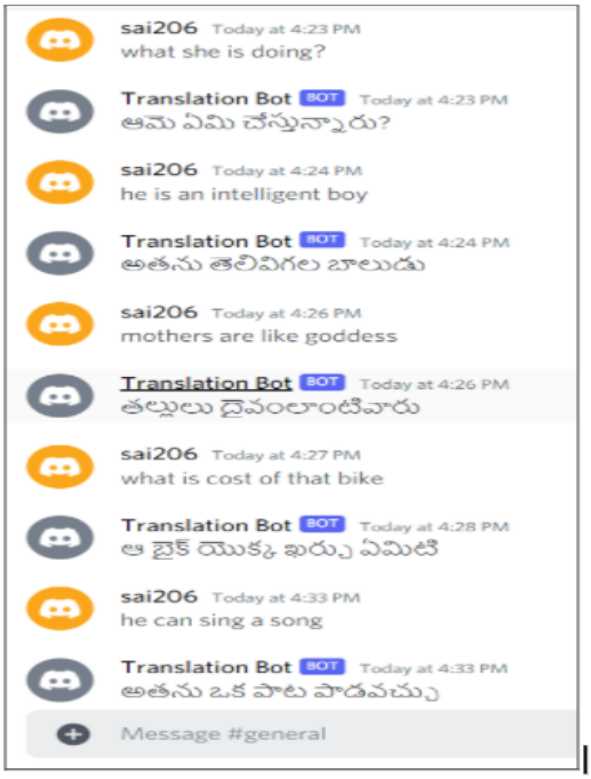

Fig. 8. Translation Bot Results

REFERENCES

[1] Ilya Sutskever, Oriol Vinyals and Quoc V. Le: “Sequence to Sequence Learning with Neural Networks”. (2014)

[2] Tiwari, G., Sharma, A., Sahotra, A., Kapoor, R. (2020). English-Hindi Neural Machine Translation-LSTM Seq2Seq and ConvS2S. 2020 International Conference on Communication and Signal Processing (ICCSP). doi:10.1109/iccsp48568.2020.91821

[3] Dzmitry Bahdanau, KyungHyun Cho and Yoshua Bengio: “Neural Machine Translation By Jointly Learning to Align and Translate” (2016)

[4] Vaswani, A., Shazeer, N.M., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I. (2017). ”Attention is All you Need.” ArXiv, abs/1706.03762.

[5] Ryan David Cunningham: “Model Compression for Chinese-English Neural Machine Translation” (2021)

[6] T. J. Sefara, S. G. Zwane, N. Gama, H. Sibisi, P. N. Senoamadi and V. Marivate, ”Transformer-based Machine Translation for Low-resourced Languages embedded with Language Identification,” 2021 Conference on Information Communications Technology and Society (ICTAS), 2021, pp. 127-132, doi: 10.1109/ICTAS50802.2021.9394996.

[7] S. Gogineni, G. Suryanarayana and S. K. Surendran, ”An Effective Neural Machine Translation for English to Hindi Language,” 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 209-214, doi: 10.1109/ICOSEC49089.2020.9215347

[8] R. F. Gibadullin, M. Y. Perukhin and A. V. Ilin, ”Speech Recognition and Machine Translation Using Neural Networks,” 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2021, pp. 398-403, doi: 10.1109/ICIEAM51226.2021.9446474.

[9] Github : https://github.com/facebookresearch/WavAugment.git

[10] Github : https://github.com/joshua-decoder/indian-parallel-corpora

[11] https://youtu.be/SPTfmiYiuok