A Structured Transformer Neural Machine Translation on Abstractive Text Summarization for Bangla seq2seq Learning

Sheikh Abujar, Akm Shahariar Azad

Department of Computer Science, The University of Alabama at Birmingham

Abstract

To the best of our knowledge, there is no universal approach that is agnostic to language despite the fact that there has been a lot of study on text summarizing for many languages. Thereby, we propose a LSTM based solution that process text in Bangla input from valuable sources after training the model the pair of Bangla text-to-Bangla summaries visualize a momentous margin that will also employ next future work.

Introduction

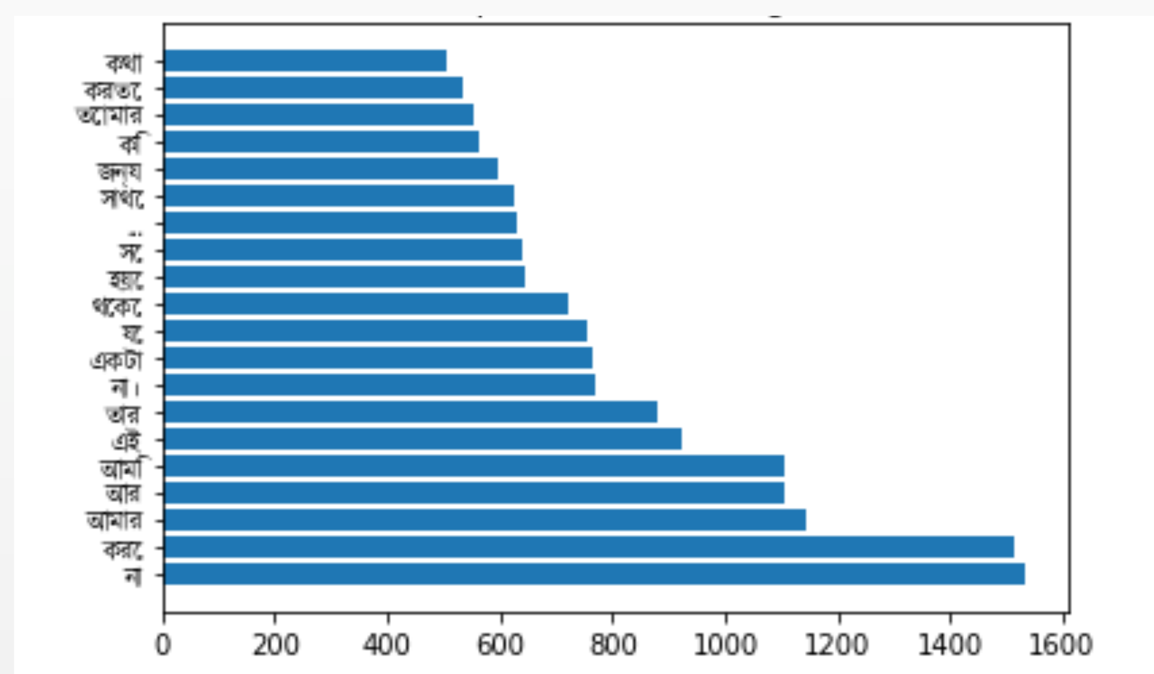
- The summary text refers to a process of language conversion in which languages are changed as a result of the efforts of neural networks.
- The ability of NMT to train a corpus into vector output in an end-to-end process allows it to translate the longest sequence of input in Bangla.
- The benefit of natural language processing's abstractive summarization of multiple languages comprises basic architecture updated by RNNs and Convolutional Neural Network encoding (CNN).

Challenge and Contribution

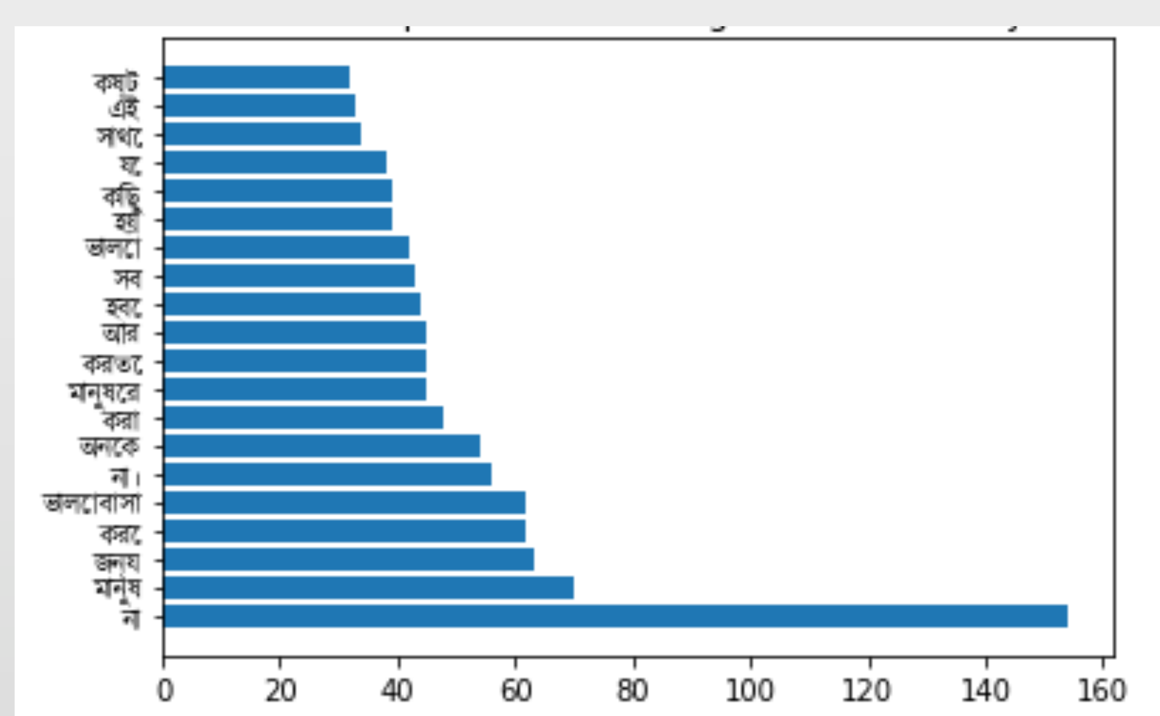
- Used self-attention which generates longer sequence translation for Bangla text input and tried best to decode the text summarization.
- To get a better enhancement and for solving difficulty in parallelization using the transformer model and getting a result of maximum length from the previous work.

Dataset properties & Statistics

- Construction in the description of the summary and text.
- Most Frequent Word in Bengali Text

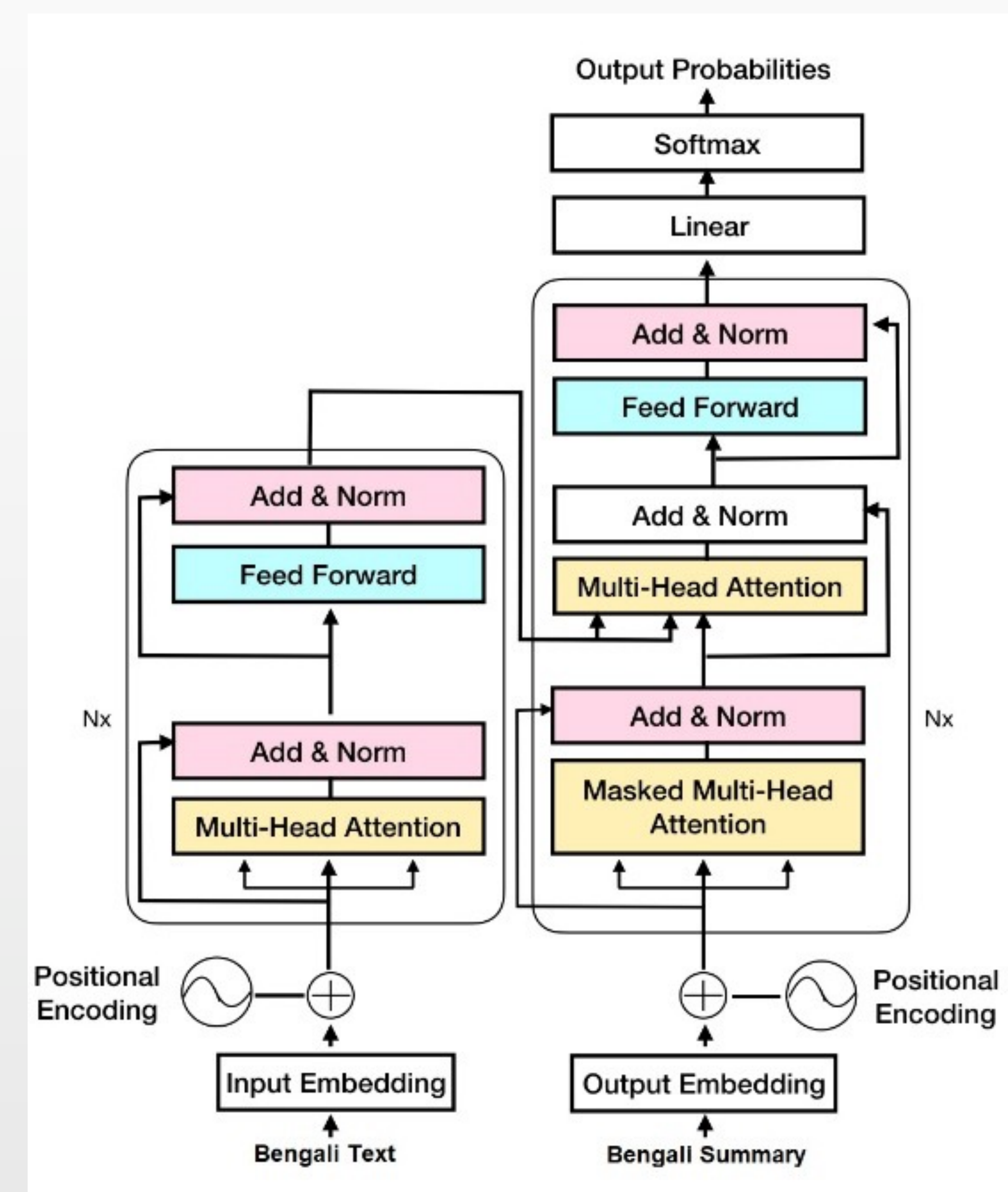


- Cleaning the texts, deleting the unwanted characters, encoding in the appropriate format, tokenized, PAD sequence, and removing the stop word.
- Most Frequent Word in Bengali Text Summary



Methodology

- Self-attention generates longer sequence translation where input Bangla text to decode the text summarization
- Neural Network Transformer defined with a model by encoding decoding with Layer.
- Multi-head attention used to 8 parallel attention levels.
- Feed-Forward Networks consist of a deeper transformation with the Relu activation function.
- Particularly encoder set up with two sub-layers of Multi head Attention for mapping the word head by the attention of SoftMax function



Results

- Attention-based transformers carry self-attention, and multi-head accomplishment and concatenate it simultaneously.
- The SoftMax function distributes the transpose function the weight of attention from the source location to the encoder and decoder.
- **Performance summary in terms of BLEU aggregate**

BLEU	
Score	0.82

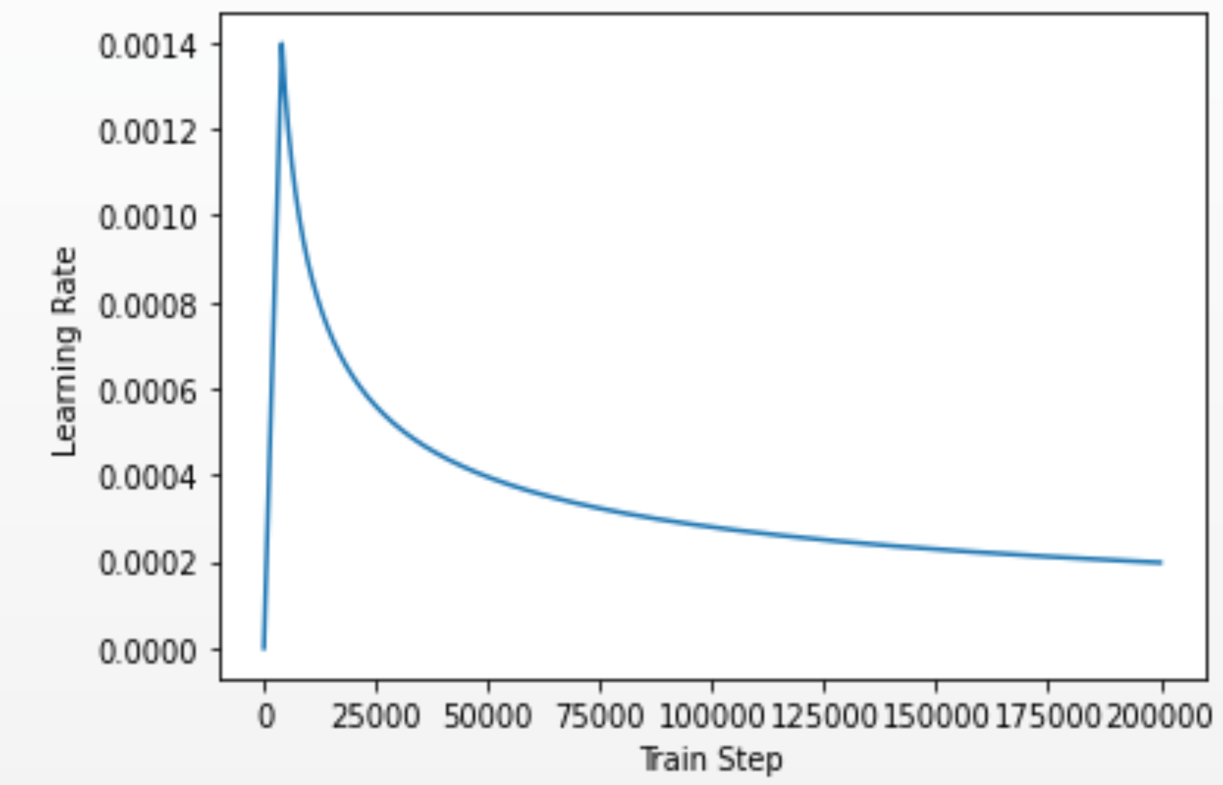
- **Rouge Score**

Rouge Score	
Score	F1-Score
Rouge 1	0.85
Rouge 2	0.83
Rouge 3	0.83

- **Transform model acceptance**

Batch Size	RNN Size	Probability	Learning Rate	EPOCHS
2	256	0.75	0.001	70
32	512	128	0.0014	100

- **Transformer model learning point scale.**



Conclusion

- Structured seq2seq process attention mechanism to explicit the Bangla language to Bangla text summarization.
- While input file analyzing the number of files was 1026 Bangla summary such as input as NMT.
- Will explore long document summarization by dynamically taking input into delivery as a response summary.

Acknowledgements

[1] A. Waibel, A. N. Jain, A. E. McNair, H. Saito, A. Hauptmann, and J. Tebelskis, "Janus: A speech-to-speech translation system using connectionist and symbolic processing strategies", in Proc. of the ICASSP, 1991, pp. 793–796.

[2] O. Kuchaiev, B. Ginsburg, I. Gitman, V. Lavrukhin, J. Li, H. Nguyen, C. Case, and P. Micikevicius, "Mixed-precision training for NLP and speech recognition with open seq2seq", 2018

[3] S. Jean, O. Firat, K. Cho, R. Memisevic, and Y. Bengio, "Montreal neural machine translation systems for wmt-15," in Proc. of the 10th WSMT. Lisbon, Portugal: Association for Computational Linguistics, September 2015, pp. 134–140. [Online]. Available: http://aclweb.org/anthology/W15-3014.

[4] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.

[5] N. Kalchbrenner and P. Blunsom, "Recurrent continuous translation models," in Proc. of the EMNLP, 2013, pp. 1700–1709.