Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 2): pp. 3110-3116
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.02.328

Deep Learning Model For Sequential Data -Machine Language Translation

Prakash Srivastava[1], Prem Ranjan Pattanayak[1], Ms.Shivani Arora[2]

[1]Department of Computer Science & Engineering, Graphic Era (Deemed to be University), Dehradun.

[2]Asst.Professor, PDP, Graphic Era Hill University, Dehradun, Uttarakhand (U.K)

ABSTRACT

This paper discusses Deep Neural Networks (DNN) and deep learning as it relates to machine translation, form of natural language processing. DNN is now a key component of machine learning methodologies. One of the best techniques for machine learning is the recursive recurrent neural network (R2NN). Recursive and recurrent neural networks are combined to create it (such as Recursive auto encoder). In this research, semi-supervised learning techniques are used to train the LSTM for reordering from source to target language. To create word vectors for the source language, the Seq2word tool is necessary, and the auto encoder aids in the reconstruction of the vectors for the destination language in a tree structure. The output of seq2word is crucial for the input vectors' word alignment. Due to the LSTM structure's complexity and time required to train the enormous data set using seq2seq. The performance accuracy is anlayzed using BLEU score.

Keywords: DNN, LSTM, deep learning model, Machine language translation

1. INTRODUCTION

In our contemporary period of globalization, where nations are interconnected in some form, language translation plays a very important function. Despite the fact that English is the most widely used language for communication, many nations still choose to use their own local tongue. If we put the global perspective away for a moment, there are still individuals who comprehend information more effectively in their own tongue. Language translation has therefore been a component of human civilization for a very long time, taking into account all of these needs. Prior to the development of technology, humans used to serve as translators between various ethnic groups. But as soon as technology advanced, we built machines to perform this translation.

The study of how software is applied to translate text or speech machine translation, the conversion of one language into another and it is a subfield of computational linguistics. Simple mechanical word substitutions are performed by machine translation, however this rarely results in accurate translation because understanding full phrases and their closest analogues in the target language is required. It's possible for two given languages to have entirely distinct structures. There are no translations for certain words in other languages.

Many words also have many meanings. The use of neural approaches [2, 7, 9] to address this issue improves translations and takes into account variances in idiomatic and typological translation. As more time goes by and experts continue to study how to make machine translation more accurate and right, this machine translation continues to evolve. Building a basic language translator is not that difficult anymore due to the advancement of deep learning and neural networks [8,10] like RNNs, LSTM, and other sequence-based neural nets.

1.1 Background

The popular tool word2vec is available to generate the vector [1]. Various models (CBOW, Skip-gram) and algorithms (Hierarchical softmax, negative sampling) work behind in word2vec processing [3]. Word2vec reduce the dimensionality of word using dimension reduction technique.

Stanford Tagger and Parser are used to analyse grammatical structures [4]. The developed module can translate a straightforward English statement. For around 500 test words, the system generated evaluation scores of 0.604 for n-gram blue, 0.830 for METEOR and F-Score was 0.816.

The module described in this study uses heuristic word information from the bilingual ANN-based word mapping module in conjunction with sentence attributes from the grammar and sentence structure analysis module to add syntax to the target language words and to flag cases to produce intelligible translation.

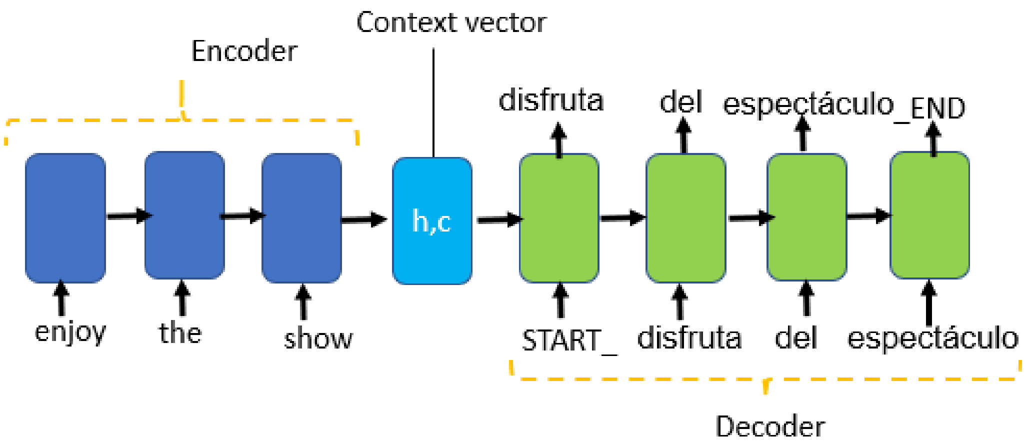
A neural network approach was employed for automatic machine translation and natural language processing in the English-Sanskrit MT system [5]. This algorithm is used for training ANN bilingual dictionary items and ANN grammatical structure.

BLEU[6] is an IBM-developed metric that contrasts the candidate translation with reference translations using modified n-gram precision. The BLEU score is calculated by multiplying the the geometric mean of the precision scores modified by the exponential brevity penalty for the test corpus.

2. DESIGN OF EXPERIMENT/ MATERIAL METHODS

2.1 Encoder-Decoder Model

The Long Short Term Memory is used by seq2seq model, usually referred to as the encoder-decoder model, to generate text from the training corpus. Additionally important in machine translation applications is the seq2seq paradigm. It predicts the possibility that a word from the user's input would appear is used to predict each of the subsequent words..

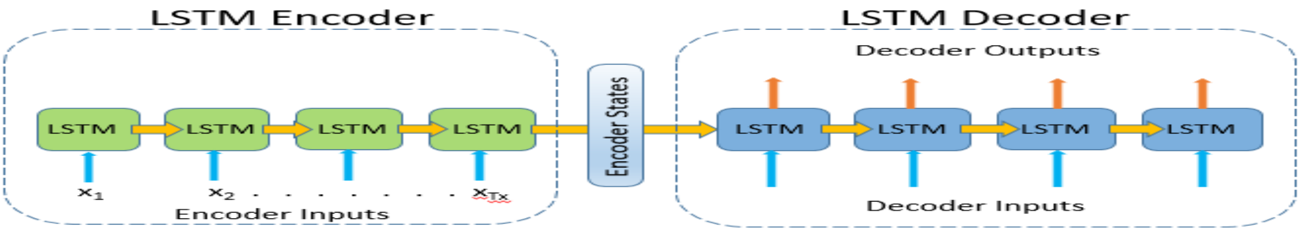


A final state vector (memory) generated by the encoder serves as the initial state of the decoder. We train the decoder using an instructional forcing technique such that it can predict the following words in a goal sequence that is given by the preceding words. To tell our decoder when to stop, we are inserting the token "<END>".

DATASET:

The dataset used in this project is an English-French language dataset containing more than 190K characters which I have taken from http://www.manythings.org/anki/. This dataset contains small sentences with a max sequence length of 16 (for input). The target sentence can contain more or less sequences than their respective input

2.2 Architecture of Sequential Data model

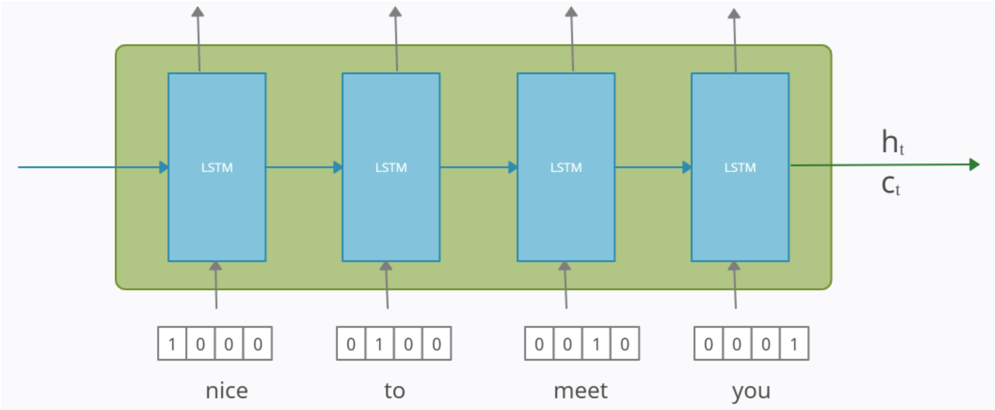


2.3 Steps in Natural Language Translation

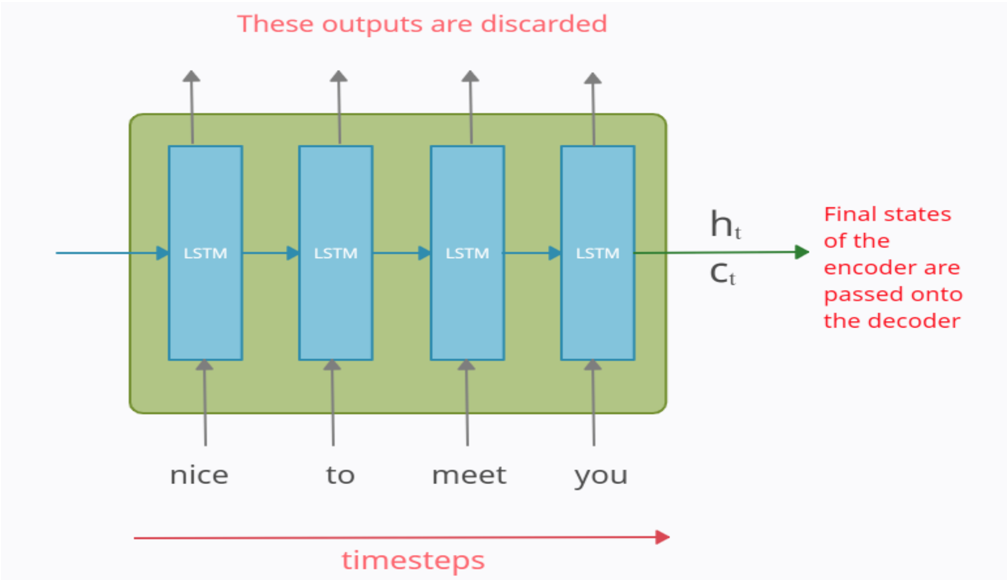
The implementation of the Lane detection model is done using the following steps:

<u>Pre-processing Input</u>: Pre-processing of input dataset by removing '\n' or '\t' is done, after that we need to represent the dataset in such form that can be taken as input by encoder i.e., in form of vectors using Embedding layer, Word2Vec or One Hot Encoding.

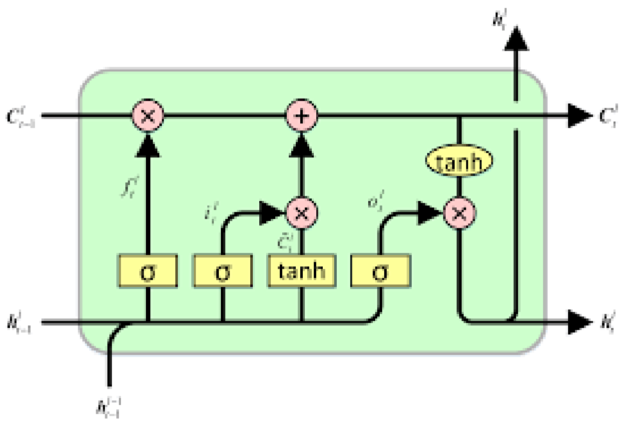
.



<u>Neural Machine Translation:</u> In order to create statistical models with the ultimate goal of translation, machine translation (MT) uses neural network models, which are based on the human brain. RNN, LSTM, and GRU, any of these neural networks could be used for our problem statement. Out of these LSTM works best because of its property of memory. CNN can be used if the input sequence was an image or video.

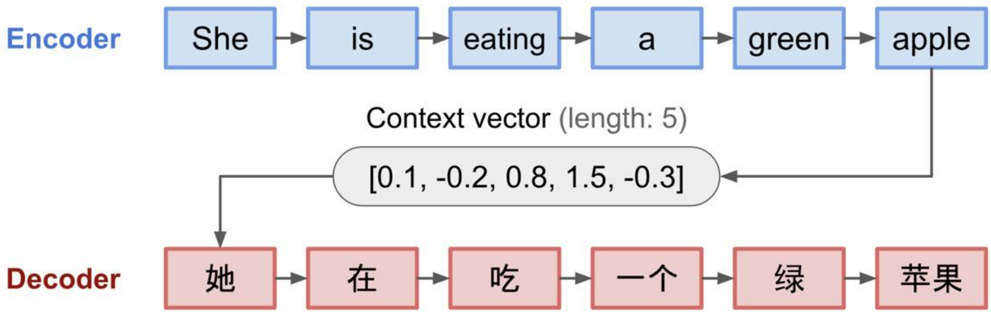


(Encoder Architecture)

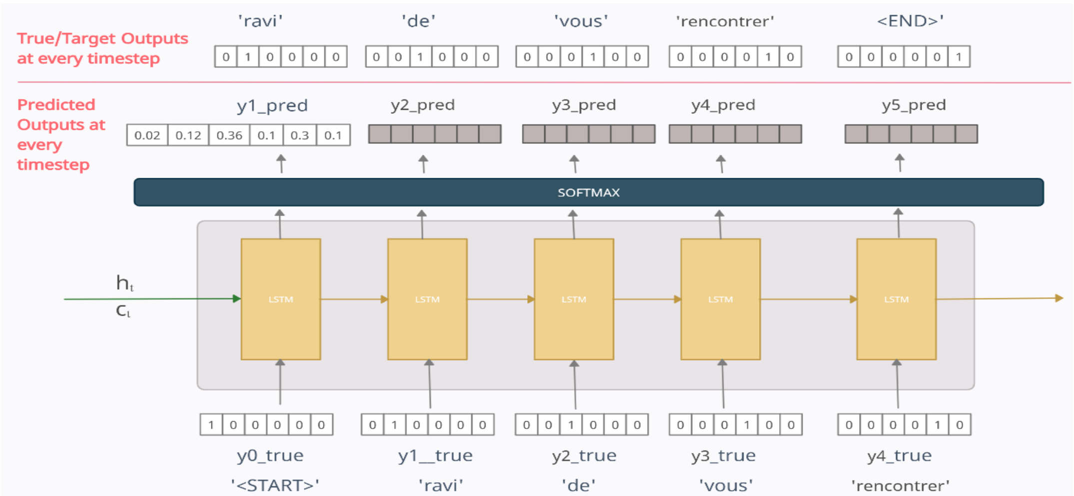


(LSTM Architecture)

Context Vector: The first cell of the decoder network receives the context vector from the final encoder cell as input. The decoder begins producing the output sequence using these initial states, and future forecasts also take into account these outputs. Here, the context vector is defined as the encoder hidden state and the cell state. The encoder output state is discarded as it is not used.

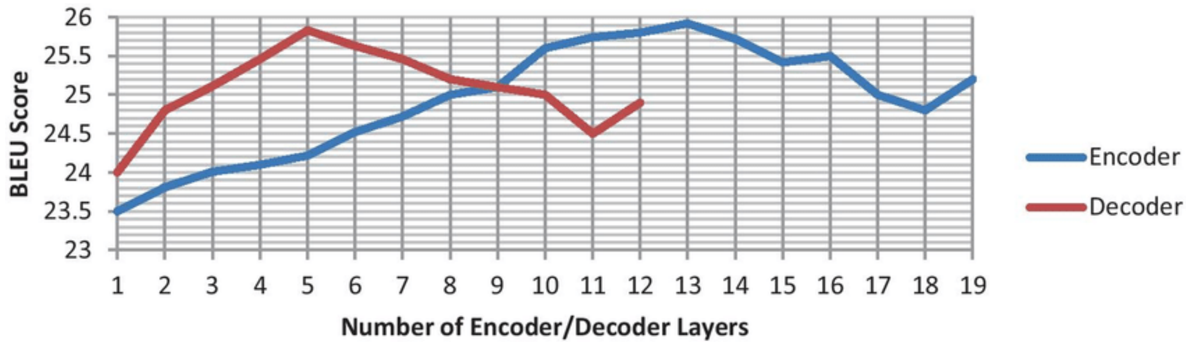


- **Decoder Output:** A collection of several recurrent units, each of which forecasts an output at time step t, y t. We compute the outputs using the hidden state at the current time step and the appropriate weight W. (S). To create a probability vector that will allow us to forecast the result, we shall use Softmax.



3. RESULTS AND DISCUSSION

This model was by far the best model for sequential data at that time. A BLEU score of 34.81 was achieved by this model. The problem was that it does not perform well when used for longer sentences. As the input and output sequences are different lengths, each batch must contain input and output phrases that are similar in length. This makes the batches tiny and renders parallelization ineffective. By first grouping phrases by the length of the encoder inputs and then for each encoder, grouping by the length of the decoder inputs, some batching is actually practicable. However, this is incredibly wasteful because we need a huge quantity of words to ensure that every set of encoder/decoder pairs is big enough to fully utilise the computing power at every training batch. It's highly unlikely but not impossible.



4. Conclusions

In the present time, machine translation is an emerging topic for research in the area of natural language processing. To train a translation system that works like the human brain, we used deep learning techniques. In contrast to other neural networks, RNN, RAE offers higher results in text processing. A well-trained deep neural network can be used for word alignment, reordering, and language modelling. RNN has the ability to apply sentence-reordering rules. GPU provides extremely parallel computation, which helps to tackle the issue of complex computation and improve system performance.

The future research will result in the creation of an optimal learning system based on quantum computing.

References

[1] Zhang, J., & Zong, C. (2017). Deep neural network in machine translation. Institute of Automation, Chinese Academy of Sciences.

[2] Sennrich, R., Haddow, B., & Birch, A. (2016). Edinburgh neural machine translation systems for WMT 16. arXiv preprint arXiv:1606.02891.

[3] Bae, S., & Yi, Y. (2016, October). Acceleration of word2vec using gpus. In International Conference on Neural Information Processing (pp. 269-279). Springer, Cham.

[4] Khan, S., & Mishra, R. B. (2012). A neural network based approach for English to Hindi machine translation.

[5] Mishra, V., & Mishra, R. B. (2010). ANN and Rule based model for English to Sanskrit Machine Translation. INFOCOMP Journal of Computer Science, 9(1), 80-89.

[6] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting of the Association for Computational Linguistics (pp.311-318).

[7] Liu, S., Yang, N., Li, M., & Zhou, M. (2014, June). A recursive recurrent neural network for statistical machine translation. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1491-1500).

[8] Bahrampour, S., Ramakrishnan, N., Schott, L., & Shah, M. (2015). Comparative study of deep learning software frameworks. arXiv preprint arXiv:1511.06435.

[9] He, W., He, Z., Wu, H., & Wang, H. (2016, February). Improved neural machine translation with SMT features. In Thirtieth AAAI conference on artificial intelligence.

[10] Eason, G., Noble, B., & Sneddon, I. N. (1955). On certain integrals of Lipschitz-Hankel type involving products of Bessel functions. Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 247(935), 529-551.