© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Machine Translation Using Deep Learning : A Survey

Janhavi R. Chaudhary[1], Ankit C. Patel[2]

[1]M. E. Student (Information Technology Department), L. D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

[2]Assistant Professor (Information Technology Department), L. D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Machine Translation using Deep Learning (Neural Machine Translation) is a newly proposed approach to machine translation. The term Machine Translation is used in the sense of translation of one language to another, with no human improvement. It can also be referred to as automated translation. Unlike the traditional statistical machine translation, the neural machine translation aims at building a single neural network that can be jointly tuned to maximize the translation performance. This survey reveals the information about Deep Neural Network (DNN) and concept of deep learning in field of natural language processing i.e. machine translation. It is better to use Recurrent Neural Network(RNN) in Machine Translation. This paper studies various techniques used to train RNN for various language corpuses. RNN structure is very complicated and to train a large corpus is also a time-consuming task. Hence, a powerful hardware support (Graphics Processing Unit) is required. GPU improves the system performance by decreasing training time period.

Keywords: Machine Translation, Neural Networks, Neural Machine Translation, Deep Learning

I. INTRODUCTION

A. Deep Learning

Deep learning is part of machine learning methods based on learning data representations, as opposed to task-specific algorithms[9]. Learning can be supervised, partially supervised or unsupervised. Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks[1] have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation and bioinformatics where they produced results comparable to and in some cases superior human experts.

B. Deep Neural Networks

A deep neural network (DNN) is an ANN with multiple hidden layers between the input and output layers[1]. DNNs are typically feed forward networks in which data flows from the input layer to the output layer without looping back. Recurrent neural networks (RNNs), in which data can flow in any direction, are used for applications such as language modeling. Long short-term memory is particularly effective for this use. Convolutional deep neural networks (CNNs) are used in computer vision. CNNs also have been applied to acoustic modeling for automatic speech recognition (ASR)[1].

C. Machine Translation

Machine Translation (MT) is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitution of words in one natural language for words in another. Machine Translation system are needed to translate literary works which from any language into native languages. The literary work is fed to the MT system and translation is done. Such MT systems can break the language barriers by making available work rich sources of literature available to people across the world[11]. Figure 1[11] shows process of Machine Translation in the form of pyramid.

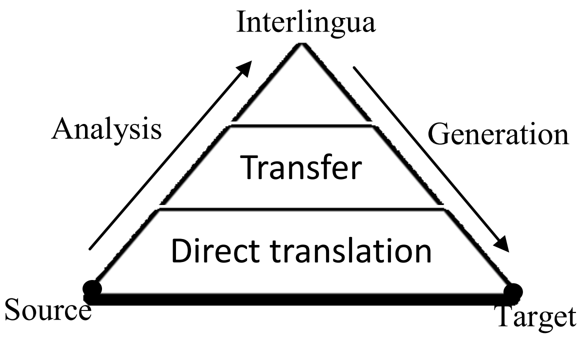

Figure 1. Machine Translation Pyramid

D. Neural Machine Translation

Neural Machine Translation is the approach of modeling the entire MT process via one big artificial neural network[18]. Figure 2[18] shows architecture or encoder and decoder.

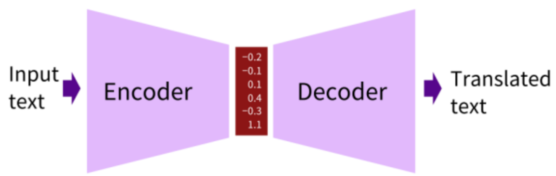

Figure 2. Neural encoder-decoder architecture

II. LITERATURE SURVEY

Jiajun Zhang and Chengqing Zong have designed many kinds of Deep Neural Networks. There are five popular neural networs introduced in their research work[1].

- ✓ Feed Forward neural network
- ✓ Recurrent neural network
- ✓ Recursive auto-encoder
- ✓ Recursive neural network
- ✓ Convolutional neural network

FNN(Feed Forward Network) :

The feed-forward neural network (FNN) is one of the simplest multilayer networks. Figure 3[1] shows an FNN architecture with hidden layers as well as input and output layers[1].

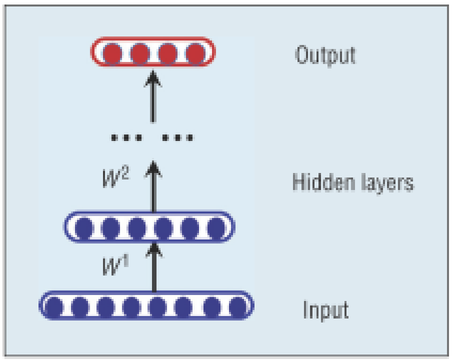

Figure 3. FNN Architecture

n-gram is collected from a text or speech corpus. The FNN attempts to predict the conditional probability of the next word given the fixed-window history words.

RNN (Recurrent Neural Network):

The recurrent neural network (RecurrentNN) is theoretically more powerful than FNN in language modeling due to its capability of representing all the history words rather than a fixed-length context as in FNN.

RAE (Recursive Auto Encoder):

The RAE provides a good way to embed a phrase or a sentence in continuous space with an unsupervised or semisupervised method.

Recursive Neural Network:

RecursiveNN differs from RAE in four points: RecursiveNN is optimized with supervised learning; the tree structure is usually fixed before training; RecursiveNN doesn’t have to reconstruct the inputs; and different matrices can be used at different nodes.

Convolutional Neural Network:

The convolutional neural network (CNN) consists of the convolution and pooling layers and provides a standard architecture that maps variable-length sentences into fixed-size distributed vectors. The CNN model takes as input the sequence of word embeddings, summarizes the sentence meaning by convolving the sliding window and pooling the saliency through the sentence, and yields the fixed-length distributed vector with other layers, such as dropout and fully connected layers.

Kyunghyun Cho, Aaron Courville, and Yoshua Bengio describe Systems that learn to attend to different places in the input, for each element of the output, for a variety of tasks: machine translation, image caption generation, video clip description, and speech recognition[2]. The attention-based neural machine translation uses a bidirectional recurrent neural network (BiRNN) as an encoder. The forward network reads the input sentence from the first word to the last, resulting in a sequence of state vectors

$$\{\overrightarrow{h_1}, \overrightarrow{h_2}, .., \overrightarrow{h_T}\}$$

The backward network, on the other hand, reads the input sentence in the reverse order, resulting in

$$\{\overrightarrow{h_T}, \overrightarrow{h_{T-1}}, .., \overrightarrow{h_1}\}$$

The use of the BiRNN is crucial if the content-based attention mechanism is used. The content-based attention mechanism relies solely on a so-called content-based scoring, and without the context information from the whole sentence, words that appear multiple times in a source sentence cannot be distinguished by the attention model[2].

Attention based Neural Network[2]

The content-based attention mechanism computes the relevance of each spatial, temporal or spatio-temporally localized region of the input,while the location-based one directly returns to which region the model needs to attend,often in the form of the coordinate such as the – coordinate of an input image or the offset from the current coordinate.

Sanjanaashree P and Anand Kumar M presented a new area of Machine Learning approach termed as a Deep Learning for improving the bilingual machine transliteration task for Tamil and English languages with limited corpus[3]. This technique precedes Artificial Intelligence. The system is built on Deep Belief Network (DBN), a generative graphical model, which has been proved to work well with other Machine Learning problem. They have obtained 79.46% accuracy for English to Tamil transliteration task and 78.4 % for Tamil to English transliteration. Bilingual Machine Transliteration task for Tamil and English languages is proceeded using DBN[3]. Figure 4[3] shows DBN architecture.

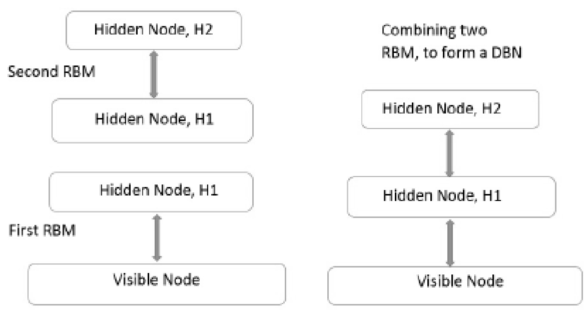

Figure 4. DBN Architecture

The system architecture is shown in Figure 5[3]. The two layers RBM on the right side is the encoders for source language and the left side is named as target language encoders. The uppermost layer is called Joint layer that concatenates the output of the top layer of source and target encoders. For a word to be transliterated, the source language word is passed through the source encoders and then reaches the joint layer and at last traverses downwards through the target encoders resulting with the transliterated word as the final output. In this architecture, each layer has 'n' number of neurons. Greedy layer-wise training is performed, the output of a layer becomes an input for the preceding layer.

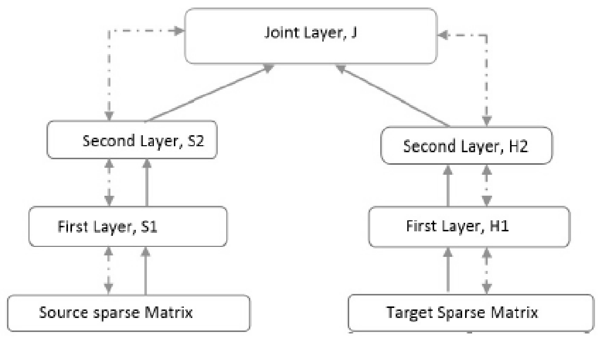

Figure 5. System Architecture for Transliterian

It concludes that a Deep Learning approach for bi-lingual transliteration is handled for English and Tamil languages using Deep Belief Networks. DBN is fully bi-directional, supports dimensionality reduction and it supports unsupervised way of learning feature and a supervised way for transliteration. with the small amount of corpus accounting to 3000 names, a decent result of accuracy around 79% is obtained in both ways, bi-lingual transliteration.

Zhen Yang, Wei Chen, Feng Wang and Bo Xu validate the hypothesis and propose a simple and flexible framework, which enables the NMT model to only focus on the relevant sense type of the input word in current context[4]. Firstly, this is the first effort to introduce the multi-sense representation, which represents each sense type of the word with a sense-specific embedding, into NMT. Secondly, propose a sense search module which can detect the sense type of the word automatically. Multi-Sense model is able to detect the sense type of the word exactly, and achieves remarkable improvements on every test set over competitive baselines. The proposed sense search module enables the model to detect the right sense type of the input word automatically. Since the sense search module is task independent, it can be applied to any other semantic related NLP tasks with little modification[4].

XIAO-XUE WANG, CONG-HUI ZHU, SHENG LI, TIE-JUN ZHAO, DE-QUAN ZHENG have proposed trilingual NMT[5]. Based on the Encoder-Decoder and attentional mechanism, we translate source language to target language, meanwhile translate another parallel source language to target language. They provide two approaches called splicing-model and similarity-model. Both of the approaches are in order to enhance the semantic representation of input sequences.

Splicing model[5]

In this model, we get a new vector c which includes information of vector c' from source language1 and vector c'' from source language2. We think the new vector c is the semantic representation of parallel source language1 and source language2.

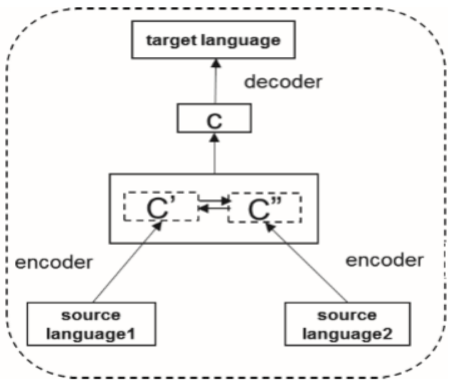

Figure 6[5]. The process of Splicing model

Similarity-model

Figure 7[5] shows the process. Because the whole process is based on similarity of vectors, we call the model similarity-model in this paper. And as you see in Figure 2, two systems are independent, so parameters are independent. Only in the training process we need to input source language1 and source language2 simultaneously. Once the model is trained, we could test the performance of the system from source language1 to target language by only inputting source language1 and test the performance of the system from source language2 to target language by only inputting source language2.

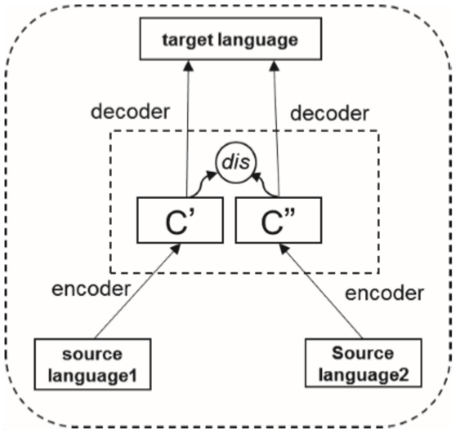

Figure 7. The process of similarity model

Feng Wang, Wei Chen, Zhen Yang, Xiaowei Zhang, Shuang xu and Bo Xu have proposed NMT model with class-specific copy network, which is referred to as CSCNMT[6]. With the network, the proposed NMT model is able to decide which class the words in the target belong to and which class in the source should be copied to. Experimental results on Chinese-English translation tasks show that the proposed model outperforms the traditional NMT model with a large margin especially for sentences containing the rare words. Copy mechanism has been proposed to deal with rare and unseen words for the neural network model using attention. But the negative point is that it's only able to decide whether to copy or not. It is unable to detect which class should the rare word be copied to, such as person, location, and organization. This paper proposes a new NMT model by novelly incorporating a class-specific copy network to overcome this issue[6].

Andi Hermanto, Teguh Bharata Adji, Noor Akhmad Setiawa have proposed RNN model for English-Indonesian machine translation[7]. In this research, a comparison between neural based network that adopts Recurrent Neural Network (RNN) and statistical based network with n-gram model for two-way English-Indonesian Machine Translation (MT) is conducted. The perplexity value evaluation of both models show that the use of RNN obtains a more excellent result. Meanwhile, Bilingual Evaluation Understudy (BLEU) and Rank-based Intuitive Bilingual Evaluation Score (RIBES) values increase by 1.1 and 1.6 higher than the results obtained using statistical based.

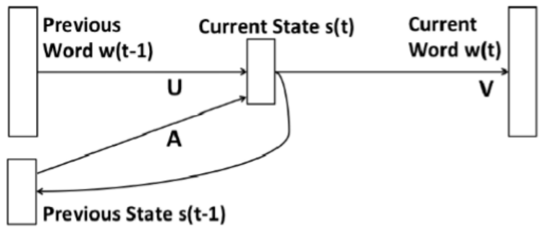

Figure 8. Recurrent Neural Language Model

Table 1. Literature Review

Review Paper	Title	Publication and year	Description
Paper-1 [1]	Deep neural networks in machine translation : An overview	IEEE, 2015	Various Neural Networks used in Different NLP methods are described
Paper-2 [2]	Describing multimedia content using attention based Encoder-Decoder network	IEEE Transactions on Multimedia, 2015	Describes Attention Based Neural Networks
Paper-3 [3]	Joint Layer based Deep Learning Framework for Bilingual Machine Transliteration	IEEE, 2015	Use of Deep Belief Network for Transliteration task for Tamil and English languages
Paper-4 [4]	Multi-Sense Based Neural Machine Translation	IEEE, 2017	Shows comparison between RNN and Sense-based model and introduces the multi-sense representation
Paper-5 [5]	Neural Machine Translation Research Based On The Semantic Vector Of The Tri-lingual Parallel Corpus	IEEE, 2016	Describes two approaches for trilingual translation with parallel corpus based on RNN and attention mechanism
Paper-6 [6]	A Class-specific Copy Network for Handling the Rare Word Problem in Neural Machine Translation	IEEE, 2017	Proposes a class-specific copy network model to overcome some issues faced in handling rare words
Paper-7 [7]	Recurrent Neural Network Language Model for English-Indonesian Machine Translation: Experimental Study	IEEE, 2015	Uses RNN language model for English-Indonesian Translation and shows comparison of this model with Statistical based language model

III. PERFORMANCE IMPROVEMENT USING GPUs

Deep learning application requires high computations because there exists large matrix multiplication, parallel processing and number of calculations during training phase. Graphics processing unit (GPU) is very good option for parallel processing and fast computation as compare to the CPU. GPU not only provides better energy efficiency but it also archives substantially higher performance over CPUs.

IV. CONCLUSION

In the present time, machine translation is a very hot research topic in natural language processing area. Deep learning helps to train a translation system like a human brain. RNN and RAE provides better result in text processing as compare to other neural networks. From above all papers about Machine Translation I have analyzed that deep learning is much better than any other methods for giving accurate result. But the processing time is high, also it needs more time to training a system.

V. REFERENCES

[1]. Jiajun Zhang and Chengqing Zong,"Deep neural networks in machine translation : An overview", in IEEE intelligent system, 2015.

[2]. Kyunghyun Cho, Aaron Courville, and Yoshua Bengio, "Describing multimedia content using attention based Encoder-Decoder network", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 17, NO. 11, NOVEMBER 2015.

[3]. Sanjanaashree P, and Anand Kumar M, "Joint Layer based Deep Learning Framework for Bilingual Machine Transliteration", IEEE, 2015.

[4]. Zhen Yang, Wei Chen, Feng Wang, Bo Xu, "Multi-Sense Based Neural Machine Translation",in IEEE, 2017.

[5]. XIAO-XUE WANG, CONG-HUI ZHU, SHENG LI, TIE-JUN ZHAO, DE-QUAN ZHENG, "Neural Machine Translation Research Based On The Semantic Vector Of The Tri-lingual Parallel Corpus", in IEEE, 2016.

[6]. Feng Wang, Wei Chen[1] ,Zhen Yang, Xiaowei Zhang, Shuang xu, Bo Xu, "A Class-specific Copy Network for Handling the Rare Word Problem in Neural Machine Translation", in IEEE, 2017.

[7]. Andi Hermanto, Teguh Bharata Adji, Noor Akhmad Setiawan, "Recurrent Neural Network Language Model for English-Indonesian Machine Translation: Experimental Study", in IEEE, 2015.

[8]. Biao Zhang, Deyi Xiong, Jinsong Su, and Hong Duan, "A Context-Aware Recurrent Encoder for Neural Machine Translation", IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 25, NO. 12, DECEMBER 2017.

[9]. Deep Learning – Wikipedia, https://en.wikipedia.org/wiki/Deep_learning .

[10]. Xiao Sun, Xiaoqi Peng, Fuji Ren, Yun Xue, "Human-Machine Conversation Based on Hybrid Neural Network", in IEEE, 2017.

[11]. Sandeep Saini, Vineet Sahula, "A Survey of Machine Translation Techniques and Systems for Indian Languages", in IEEE, 2015.

[12]. Eric Greenstein, Daniel Penner, "Japanese-to-English Machine Translation Using Recurrent Neural Networks", available at https://github.com/lisagroundhog/GroundHog.

[13]. Bing Zhao, Yik-Cheung Tam, "BILINGUAL RECURREN T NEURAL NETWORKS FOR IMPROVED STATISTICAL MACHINE TRANSLATION", in IEEE, 2014.

[14]. Preeti Dubey, "Need for Hindi-Dogri Machine Translation System", in IEEE, 2014.

[15]. Heeyoul Choi, Kyunghyun Cho, Yoshua Bengio, "Context-dependent word representation for neural machine translation", in Elsevier, 2017.

[16]. Todd Law, Hidenori Itoh, Hirohisa Seki,"A neural network assisted Japanese-English Machine translation system", International Joint Conference on Neural Networks, 1993.

[17]. KyungHyun Cho, Yoshua Bengio,"NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE ", in ICLR, 2015.

[18]. Neural Machine Translation Tutorial- https://sites.google.com/site/acl16nmt/home