© 2018 IJSRSET | Volume 4 | Issue 4 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Bilingual Machine Translation Using RNN Based Deep Learning

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

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

[2]Assistant Professor(Information Technology Department), L. D. college of Engineering, 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. The models proposed recently for neural machine translation often belong to a family of encoder–decoders and encode a source sentence into a fixed-length vector from which a decoder generates a translation. In this approach Neural Machine Translation technique will be used to translate Japanese language into English language. Tanaka corpus will be used in this approach. Japanese will be translated into English using improved Recurrent Neural Network(RNN).

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.

Deep learning is one step further of Representation learning in which there are multiple levels of features. These features are automatically discovered and composed together at various levels which are mapped to output. Each level represents set of features which are discovered from the features represented in the previous level. In neural network multiple layers are corresponding to multiple levels of features. Today's advanced deep neural networks use algorithms, big data and computational power of GPUs to change the dynamics of Represenatation learning. Deep learning is used to help solve many big data problems such as computer vision, speech recognition, and natural language processing.

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.

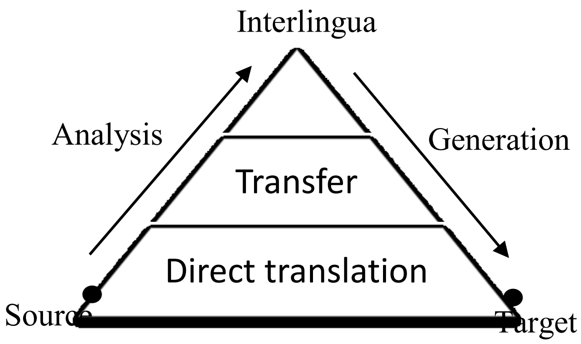

Figure 1. Machine Translation Pyramid

D. Neural Machine Translation

Neural machine translation is a newly proposed approach for machine translation using special neural network framework called Encoder-Decoder Architecture. This section describes a simple encoder-decoder architecture for neural machine translation which is originally proposed by [Cho et al.2014b].

As we discussed in deep learning dynamics, the neural machine translation doesn't require predefined features. Predefined features mean those features which are designed, not learned from the data. The goal of Neural Machine Translation is to build a fully trainable model of which every component is trained and tuned together using paraller corpora to maximize the transalation performance.

A fully trainable Neural machine translation model starts with raw representation of source sentence. Without loss of generality and for sake of simplicity let us consider the smallest unit of source sentence is word. A raw representation of source sentence is a sequence of words. Each word is encoded by one-hot vector, which is in simple terms a vector representing its integer index in a fixed size vocabulary.

II. LITERATURE SURVEY

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	IEEE, 2015	Use of Deep Belief Network for Transliterati

	for Bilingual Machine Transliteration		on 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	IEEE, 2015	Uses RNN language model for

	Language Model for English-Indonesian Machine Translation: Experimental Study		English-Indonesian Translation and shows comparison of this model with Statistical based language model

III. EXPERIMENT

Step 1 : First of all collect Japanese-English vocabulary data.

Step 2 : Create RNN for Encoding texts with help of python.

Step 3 : Train the system with languge corpus.

Step 4 : Translate Japanese language into English language.

Given figure shows Text processing in Machine translation:

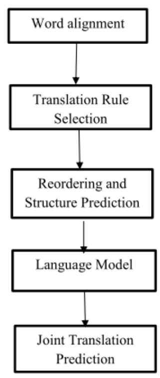

Figure 2

In basic RNN model the RNN Encoder reads the input source sentence from say x_1 to x_T where T is length of the source input sentence and compress into history/summary vector h_T .In proposed scheme the bidirectional RNN encoder is used.

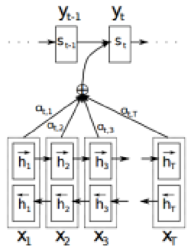

Figure 3. Bidirectional RNN Encoder

A bidirectional RNN encoder consists of the forward RNN and the backward RNN as shown in figure. The forward RNN reads the input source sentence from x_1 to x_T and computes a sequence of forward hidden states from h_1 to h_T . The backward RNN reads the input source sentence from x_T to x_1 and computes a sequence of forward hidden states from h_T to h_1 . An annotation vector h_j for each source word x_j is obtained by concatanating the h_j and h_j in order to capture the summarization not only based on preceeding words but also on the following words. Since the RNN has a tendency to represent the recent inputs well the annotation vector h j is more focused on surrounding of word x_j. In other words, each annotation vector h_i captures information about whole sentence with strong focus given on surrounding of i-th word of the input sequence[17].

IV. RESULTS AND DISCUSSION

For training the RNN encoder decoder, we have used Tanaka corpus. It contains 1,50,000 sentences pairs and it is publicly available database.

```
Input: UNK が ひろし に 試練 を UNK て いる 。
Target: the yakuza were tormenting hiroshi . <
Input:   UNK が ひろし に 試練 を UNK て いる
Output:  the yakuza were tormenting hiroshi .
```

Figure 4. Sample training set translations

```
Parsed Input: 彼女 は コンサート に 行く 。 <eos>
0.00663540713026: She is going to the concert .
5.27218884914: She is my professor .
6.86046250971: She is UNK .
9.47003285695: She is going is not my professor .
9.67765936555: She is going to the He .
Input Sequence: 彼女 は 家 に 行く 。
How many samples? 5
Parsed Input: 彼女 は 家 に 行く 。 <eos>
0.291400269567: She is going to school .
1.76811346484: She is going home .
2.53733918163: She is my friend .
6.52093583348: She is my boss .
7.40240003171: She is UNK .
```

Figure 5. Sample test set translations. The correct translations are marked in red.

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.

V. CONCLUSION AND FUTURE WORK

In this approach we have tried to get accurate translation from Japanese to English. We have used small group of data to train the system. We can apply this architecture to the large amount of data with efficient processing unit. In future we will try to translate accurately Japanese to Hindi.

VI. 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 Chen1] ,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