A Study of English-Vietnamese Machine Translation based on Deep Learning

Department of Information Technology
Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok,
Bangkok, Thailand

Department of Information Technology Management
Faculty of Information Technology and Digital Innovation
King Mongkut's University of Technology North Bangkok,
Bangkok, Thailand

***Abstract*—Machine translator (MT) is a popular program to translate text or speech from one language to another. MT also performs mechanic substitutions of words in one language for words in another but that rarely produces a good translation because recognition of whole phrases and the words sometimes belong to the context of conversation. In this paper, we build a MT for an English-Vietnamese translator using deep learning methods: Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Attention, and Transformer. The deep learning based machine translators were compared based on the test accuracy of results translation. It was found that best deep learning based machine translator model was the Attention mechanism, achieving 98.8% accuracy. The Transformer yielded second rank or 96.5% accuracy.**

Keywords—machine translator, deep learning, rnn, lstm, gru, attention, transformer

I. INTRODUCTION

With increasing many kinds of data in the world, getting useful information from those is very important to make a precise decision in learning and management. Some of data could develop strongly in the future especially image, text and speech [1]. To get the information of data to understand could help people to find a suitable choosing and reduce cost in time and money. Analyzing the text by learning deep neural network models to appreciate or translate is a new approach where we could know and have more information for expressing appropriate manner in a conversion.

There has been some researchers have done experiments for machine learning translators for translating from English to Vietnamese and Vietnamese to English. For example, reference [2] built three models: RNN, COnvS2S and Transformer. They reported that Transformer was the best results when given a suitable batch size and learning rate. However this model did not show the full comparing to get the best way for translation.

In this paper, we will build models base MT to translate language from English to Vietnamese. Five deep learning models were experimented based on RNN, LSTM, GRU, Attention mechanism, and Transformer model. The objective was to compare them to get the best solutions in English-Vietnamese translation.

II. LITERATURE REVIEWS

A. Machine Translator

Machine Translator (MT) is a kind of computational linguistics basing on application of software to translate text or speech from a language to another. In this work, MT shows the replacement the words by the words, but sometimes it is not enough good because the words are not corresponding meaning in some situation. Current MT is

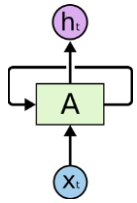

Fig. 1. Recurrent Neural Network model

improved every day and allows users to translate more and

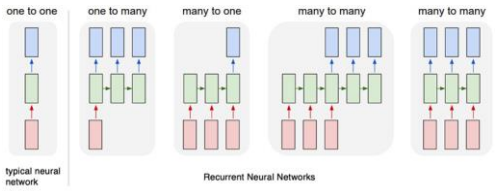

Fig. 2. Diagram of different RNN sequence types

more accurate by awareness of machine and by the methods or algorithms. In this paper, we apply the function of machine to build the models and MT will analyze by algorithms to recognize the outputs of the language from inputs.

B. Deep Learning

Deep learning is a kind of machine learning methods base on artificial neural network with representation. Deep learning is also a class of machine learning algorithms that use many layers functions to extract information from raw input data and sometimes pass over the human thinking. Deep learning could be applied to supervised, semi-supervised and unsupervised learning tasks. This is very important because some unlabeled data is good to label data. For example, deep structure could be trained in an unsupervised manner are neural history compressors and deep belief networks.

C. Recurrent Neural Network (RNN)

RNN is a kind of artificial neural networks where neuron-like nodes are organized into successive layers. Each node in layer is connected to every other node in next successive layers. RNN could use memory to process variable length sequences of input. This make them useful tasks such as unsegmented, connected handwriting or speech recognition.

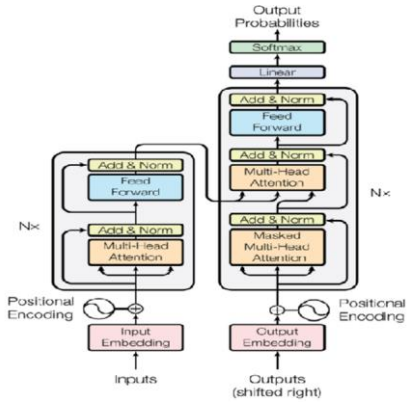

Fig. 7. Transformer model Architect

D. Long Short Term Memory (LSTM)

LSTM is a kind of artificial recurrent neutral network architecture used in deep learning. LSTM has feedback connections. It could process single data points but also entire sequences of data. For example, LSTM is applied to unsegmented, connected handwriting or speech recognition. LSTM is also good way for classification, processing and prediction based on time series data. LSTM is moreover applied for dealing with the vanishing gradient problem.

E. Gated Recurrent Unit (GRU)

GRU is a kind of artificial recurrent neutral network architecture used in deep learning. GRU combined information form a new source annotation vector to generate context vector. GRU could perform on tasks of polyphonic music modeling, speech and natural language processing similar to LSTM. GRU is shown better performance on certain smaller and less frequent datasets.

F. Attention Mechanism

Attention is a mechanism that is developed to improve the performance of Encoder-Decoder RNN on machine translation. Attention in deep learning is to confirm the factors when processing data. Attention is one component of network architecture and is in charge of managing and quantifying the relation between inputs and outputs. Attention is applied in natural language processing tasks especially in problem of long sequence in machine translation.

G. Transformer Model

Transformer model is a kind of deep learning used in natural language processing. It is also used to handle sequential data for translation and task summarization. However, unlike RNNs, Transformer does not need sequential data be processed in order. For example, in natural language, Transformer does not need to process from the beginning of it before the end so it sometimes is more parallel than RNNs so reducing training times.

III. Research Methodology

To achieve the objective, the research methodology in this research as 3 major steps: data preparation, model settings, and model training and evaluation.

1) Data Preparation

In this paper, we use a language data set retrieved from http://www.manythings.org/anki/vie-eng.zip. This data set contains three columns of sentences in English, Vietnamese, and descriptions. The data had totally … records divided into 50% training set, 25% validation set, and 25% test set.

2) Model Settings

Five deep learning models were used including RNN, LSTM, GRU, Attention and Transformer. Each model had their setting as Table 1.

TABLE I. Model Settings

Models	Type	#Layers#Cells	#Parameters
RNN	Simple RNN	5, (73,163)	260623
LSTM	LSTM classic	5,(73,163)	809891
GRU	Gated recurrent units	5,(73,163)	619427
Attention	AttentionLayer	5,(24,1771)	3832271
Transformer	BERT Pytorch	5,(24,24)	109500696

3) Model Training and Evaluation

Python programing was used to write scripts for all models. The experiments were run on Google colab research https://colab.research.google.com/. The performance of the trained models based on Accuracy, Precision, Recall, F-measure as the following formula:

IV. Experimental Results

While training different models, it was observed that LSTM based neural networks took the least time to execute a training epoch, RNN took the most time and GRU took slightly greater time than LSTM. An important factor affecting the performance of these models was training data

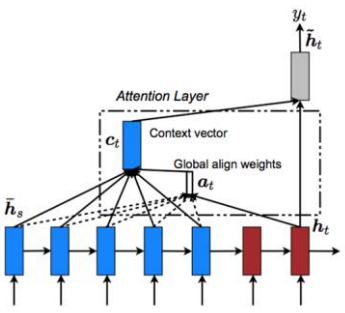

Fig. 5. Attention Model

set as there was a limited amount of structured data. In this paper, we will compare the accuracy between the model and then choose the appropriate model for translating from English to Vietnamese. In this case, we could choose Attention combining LSTM and Transformer will get the best accuracy to translate English to Vietnamese. Table 2 shows the test accuracy performance based the deep learning neural networks: RNN, LSTM, GRU, Attention, and Transformer. It shows that the Attention model is the best while Transformer is the second in ranks.

TABLE II. TABLE TYPE STYLES

Model	Loss	Val-loss	Accuracy	Time Epoch at 50 (s)
RNN	0.593	1.022	0.834	54
LSTM	0.128	0.722	0.940	132
GRU	0.148	0.625	0.937	107
Attention	0.006	0.110	0.988	10
Transformer PYTORCH	0.052	0.210	0.965	11

V. CONCLUSION AND FUTURE WORK

This paper achieves the goal of generating scripts from three different deep learning models. The models for text generation are trained using Bidirectional RNN, LSTM, Attention, Transformer and GRU. The performance of the models is further analyzed to reach a conclusion that Transformer and Attention generates text a in most efficient way followed by GRU and then SimpleRNN while loss is least in SimpleRNN followed by LSTM and it is most in GRU. The LSTM model takes the least time for text generation, GRU takes slightly more time and RNN takes the highest time. The results obtained from different models are juxtaposed through the generated graphs and also, the scripts generated are presented effectively in this paper [3][4]. Additionally, software implementation and related works in this domain are covered in different sections. We believe that further research can enhance this model and optimize it with lots of computation and data.

In the future, we will apply more models and combining them to get the best answer for translating from English to Vietnamese. With the development of technology, analytics in image and text [5] will have a new face to appreciate and adjust the behavior for a company or people in general [6]. Beside, analyzing data could help people decide perfectly in choosing a suitable method for curing to get better health [7][8]. We will apply PySpark to appreciate the nature of words relating the jobs of the questioners and get insight of the text. We will also illustrate bigger data to get more exact results. Analyzing the text data is the new directions for IT and general manger to acquire the best insight for detecting a good policy at the present and in the future [9]. For example, proposing modern condition in a park after detecting many insincere questions from users. We also will analyze the cognitive efforts users spend in writing their posts and role of that in detecting insincere content.

REFERENCES

[1] Maryam Khanian Najafabadi, Mohd Naz'ri Mahrin, Suriayati Chuprat, Haslina Md Sarkan, Improving the accuracy of collaborative filtering recommendations using clustering and association rules mining on implicit data, 2016, page 9-10.

[2] Hong-Hai Phan-Vu, Viet-Trung Tran, Van-Nam Nguyen, Hoang-Vu Dang, Phan-Thuan Do, Neural machine translation between Vietnamese and English: an empiracal study, Journal of Computer Science and Cybernetics, 2018, page 1-15.

[3] Richard A. Plunza, Yijia Zhoua, Maria Isabel Carrasco Vintimillab, Kathleen Mckeownc, Tao Yud, Laura Uguccionia, Maria Paola Sutto , Twitter sentiment in New York City parks as measure of well-being, Science Direct, 2019, page 3.

[4] Yu T, Hidey C, Rambow O, McKeown K, Leveraging sparse and dense feature combinations for sentiment classification, Science Direct, 2017, page 5,6.

[5] Maryam Khanian Najafabadi, Mohd Naz'ri Mahrin, Suriayati Chuprat, Haslina Md Sarkan, Improving the accuracy of collaborative filtering recommendations using clustering and association rules mining on implicit data, 2016, page 9-10.

[6] Chun-Wei Tsai, Chin-Feng Lai, Han-Chieh Chao and Athanasios V. Vasilakos, Big data analytics: a survey", Tsai et al. Journal of big data, 2015, page 6-7.

[7] Karthik Kambatla, Giorgos Kollias , Vipin Kumar , Ananth Grama , Trends in big data analytics", Science Direct, 2014, page 8-9.

[8] Wullianallur Raghupathi and Viju Raghupathi, Application of big data analytics in healthcare: promise and potential, Health Information Science and System, 2014, page 2-3 .

[9] Ohbyung Kwon, Namyeon Lee, Bongsik Shin, Data quality management, data usage experience and acquisition intention of big data analytics, International Journal of Information managemnet, 2014, page 4-5.