INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

Literature Survey on Natural Language Processing in LSTM- Text Generation

P Shalu, Meera M

PG Student, Dept. of C.S, GECI, KTU University, Kerala, India

Assistant Professor, Dept. of C.S, GECI, KTU University, Kerala, India

ABSTRACT: Natural language processing is always a recent area where Text generation is a subfield of a Natural language processing. Which Artificial intelligence will automatically generate the natural language texts based on the requirement. Here, we are going to study the different ways of text generations that have been developed in the last few years. Mainly focus on the evaluation of recently proposed work and neural network models. We then present the different models used and conclude the paper by proposing the future research work.

KEYWORDS: Text summarization, Natural language processing, Long Short Term Memory (LSTM), Neural Network

I. INTRODUCTION

Natural Language Processing (NLP) is a subfield of computer science,artificial intelligence which focuses on developing the system that allows computers to communicate with people using different languages . A summarization is taking a large selection of text and then reducing it to smaller pieces. Summarization consists of two types mainly Abstractive and Extractive. Extractive Summarization is to create the summary from phrases ,words, sentences etc of the input text document and Abstractive Summarization is to require deep understanding and reasoning over the text where it provides the summary without using the same word or sentences in the input text.

Deep Learning becomes advanced resulting in a variety of different research fields and application domains where we use it. In this work, we aim to provide an extensive empirical evaluation from different deep learning models for text language generation. Goal is to quantitatively assess whether LSTM-like neural networks share more similarities with natural language which humans most commonly produce.

Consequently, In this paper we focus on abstractive summarization methods and provide an overview of some of the most approaches used and some of the latest areas based on it. There are so many papers that provide abstractive overviews of text summarization techniques and systems which are proposed and get most of them a good accuracy by comparing it with other techniques.

1.1   SEQUENCE-TO-SEQUENCE MODEL

One of the best and simple Abstractive Text summarization is Sequence to sequence model. The Fig. 1, the architecture of sequence to sequence consists of two parts mainly Encoder and the Decoder.
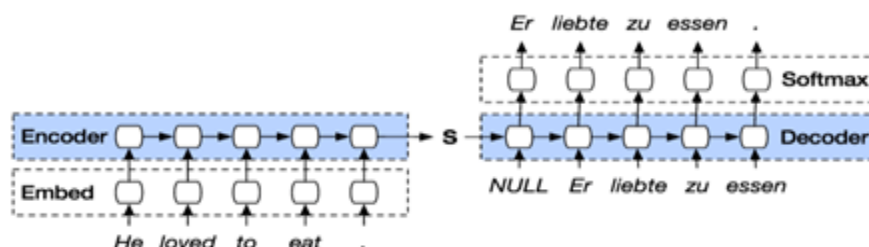

Fig. 1. Sequence to sequence Model

The Encoder is inserting the input text one word at a time then it will pass through an embedding layer that transforms the word into distributed representation. Which will combine using a multilayer neural network which contains a hidden layer generated after inserting the previous word for the first word in the text. Decoder takes as input from a hidden layer which is generated after insert in the last word of input text. The decoder will decode to generate the text summaries using an softmax and attention mechanism.

1.2 LONG SHORT TERM MEMORY

As a solution to the problem of Vanishing Gradients is LSTM. LSTM consists of common unit that composed of a cell, input gate, an forget gate and an output gate.Variant of RNNs that introduce a number of special, internal gates Internal gates help with the problem of learning relationship between both long and short sequences of data It used in speech recognition, handwriting recognition and machine translation. Fig. 2 shows the LSTM Model
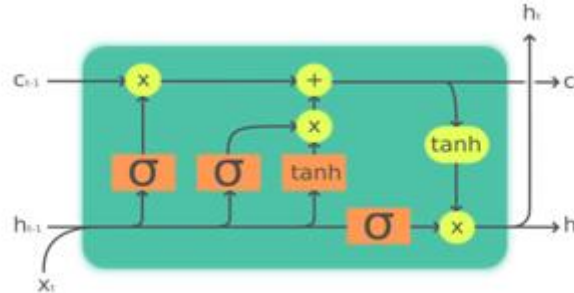

Fig. 2. Long short term Memory

1.3 RESTRICTED BOLTZMANN MACHINE

RBM (Restricted Boltzmann Machine) is an artificial neural network that learns probability distribution and it consists of visible layer and hidden layer and a neuron of each layer contains no connection between them but it is connected to each neuron of other layers. Layer in RBM form is bipartite graph

1.4 RECURRENT NEURAL NETWORK

Recurrent Neural network (RNN) is a neural network model that consists of different layers connected to each other on work based on it. The structure of the network is similar to a feedforward neural network which can handle sequential data, consider the current input , previously received inputs and also can memorize previous input due to its internal memory.

II. LITERATURE REVIEW

Konstantin Lopyrev, In this paper author introduced Generating News Headlines with Recurrent Neural Networks. Encoder-Decoder recurrent neural network with LSTM units and attention mechanism for generating the new headlines using the text. The simplified attention mechanism is used for the small set of neurons for computing the attention weights.

Shashi Pal Singh, In this paper author introduced Bilingual Automatic Text Summarization Using Unsupervised Deep Learning. Here it analyzed and compared the performances in two languages Hindi and English using an unsupervised deep learning approach. It extracts the eleven features from each sentence of the document and generating the feature matrix is passed through the Restricted Boltzmann Machine. The main limitation is it will work in multiple document summarization.

Haijun Zhang, In this paper author introduced Understanding Subtitles by Character-Level Sequence-to-Sequence Learning. The author has analyzed and compared the performances in English to Chinese subtitle translation and it's embedded an RNN into the encoder-decoder approach for generating the character level sequence representation. It can also improved by GRU in language model of encoder.

Jianpeng Cheng Mirella Lapata, In this paper author has proposed Neural Summarization by Extracting Sentences and Words where data-driven approach based on neural network and continuous sentence features. The general framework which used in single document summarization is composed of a hierarchical document encoder and the attention mechanism based on extractor.

Shengli & Haitao Huang & Tongxiao Ruan, In this paper the author introduced Abstractive text summarization using LSTM-CNN based deep learning where extractive text summarization model is concerned with syntactic structure and abstractive text summarization model are concerned with semantics. Here it work both semantics and syntactic structure in this system.

Mahmood Yousefi-Azar , Len Hamey, In this paper the author introduced Text summarization using unsupervised deep learning, the method of extractive query-oriented single document summarization using an deep auto-encoder to find feature space from the term-frequency input. The main limitation is computational cost of training.

Abu Kaisar Mohammad Masum, In this paper author introduced Abstractive method of text summarization with sequence to sequence RNNs where it translate English to English text summary using the encoder and decoder with LSTM. The main limitation is correct will be short text only, Incase long text it will be incorrect summary.

Su Zhao, Encong Deng, In this paper the author introduced Generating summary using sequence to sequence model where it deal with the problem existing in generating abstracts and get better effect of generating text abstract which can be applied to text extraction tasks. The main limitation is to add more weight parameters and also increase training time.

Shayak Chakraborty, In this paper author introduced Study of Dependency on number of LSTM units for Character based Text Generation models where it increase of LSTM cells and also increase the semantic relationship between the characters. The limitation is small corpus language character based text generation is not good. The solution is Neural network with an average number of LSTM cells.

Sandeep Saini, In this paper the author introduced Neural Machine Translation for English to Hindi require vey less amount of size for training and thus exhibit satisfactory translation for few thousand of training sentences as well. The main limitation is not good in sentences using an smaller data sets.

III. CONCLUSION

The growth of the internet in text is increasing day by day. Chance in problem, So the best solution to solved is strong text summarization which produces a summary of the document to help users in time consuming tasks. So, at this time Abstractive Text Summarization play an important role by providing a whole text document .

In this paper, we emphasized various approaches for single and multi document summarization. We already describe some of the methods that are used and algorithms. The paper provide a literature review of various research works in the field of text summarization. This research will explore existing systems and new techniques of Natural Language processing and machine learning.

REFERENCES

1. Jianpeng Cheng ,Mirella Lapata, "Neural Summarization by Extracting Sentences and Words", ACL2016 conference paper with appendix, Computation and Language (cs.CL).
2. Shayak Chakraborty; Jayanta Banik; Shubham Addhya; Debraj Chatterjee, "Study of Dependency on number of LSTM units for Character based Text Generation models", 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)
3. Abu Kaisar Mohammad Masum; Sheikh Abujar; Md Ashraful Islam Talukder; A.K.M. Shahariar Azad Rabby; Syed Akhter Hossain, "Abstractive method of text summarization with sequence to sequence RNNs", 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT),
4. Sandeep Saini, Sandeep Saini," Neural Machine Translation for English to Hindi", 2018 Fourth International Conference on Information Retrieval and Knowledge Management.
5. Shashi Pal Singh, Ajai Kumar, Abhilasha Mangal, Shikha Singhal, "Bilingual Automatic Text Summarization Using Unsupervised Deep Learning", International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016.
6. Shengli Song & Haitao Huang & Tongxiao Ruan, "Abstractive text summarization using LSTM-CNN based deep learning", Multimedia Tools and Applications volume 78, pages857–875(2019), Springer
7. Mahmood Yousefi-Azar∗ , Len Hamey, "Text summarization using unsupervised deep learning", Expert Systems with Applications, 2017 – Elsevier.
8. Haijun Zhang, Member, IEEE, Jingxuan Li, Yuzhu Ji, and Heng Yue, "Understanding Subtitles by Character-Level Sequence-to-Sequence Learning", IEEE Transactions on Industrial Informatics (Volume: 13, Issue: 2, April 2017).
9. Su Zhao, Encong Deng, Mengfan Liao, Wei Liu , Weiming Mao ," Generating summary using sequence to sequence model", 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020).

10. Marco Lippi , Marcelo A Montemurro , Mirko Degli Esposti, and Giampaolo Cristadoro, "Natural Language Statistical Features of LSTM-generated Texts", IEEE Transactions on Neural Networks and Learning Systems (Volume: 30, Issue: 11, Nov. 2019).
11. K. Greff, R. K. Srivastava, J. Koutn´ık, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Transactions on Neural Networks and Learning Systems, vol. PP, no. 99, pp. 1–11, 2017.
12. H. W. Lin and M. Tegmark, "Critical Behavior from Deep Dynamics: A Hidden Dimension in Natural Language," arXiv preprint arXiv:1606.06737, 2016.

BIOGRAPHY

P Shalu is a PG student in Computer Science and System Engineering, Government Engineering College, Idukki, KTU University,Kerala, India. I completed my Btech in Computer Science Engineering , Axis college of Engineering,Thrissur, Kerala, India. Research interests are NLP, Artificial intelligence , Machine learning etc.

Meera M is a Assistant Professor in Computer Science and Engineering, Government Engineering college, Idukki,Kerala. She received Mtech (2014) in Computer Science and InformationSystems, CUSAT,Kerala. Her research interest in Natural language processing,Machine Learning etc.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING