Conversing verses - haiku generation using a LSTM-based auto-encoder matching model

Luka Ernestini
Univerza v Mariboru
Fakulteta za elektrotehniko, računalništvo in informatiko
Maribor, Slovenija
luka.ernestini@student.um.si

Niko Uremović
Univerza v Mariboru
Fakulteta za elektrotehniko, računalništvo in informatiko
Maribor, Slovenija
niko.uremovic@um.si

ABSTRACT
Lorem ipsum

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous

General Terms
Theory

Keywords
text generation, neural networks, LSTM

1. INTRODUCTION

Deep learning is an emerging machine learning approach that can already be seen applied in many industries, including natural language processing. Example use cases include identifying people and objects in images and videos, understanding voice commands (smartphones, cars, smart houses) and providing better results for internet search queries. This technology brought machines the closest they have ever been to how we humans think and talk. Having learned the rules and grammar of our natural language, machines are now able to generate text for various applications that is in some cases indistinguishable from human written text [7].

On the field of natural language processing, many methods deep learning have been developed for text generation. Song, Huang and Ruan [9] achieved Abstractive Text Summarization (ATS) using and LSTM-CNN based ATS framework (ATSDL). You et al. [10] researched ways to generate natural language descriptions of images. They proposed a model of semantic attention which combines the established topdown and buttom-up approaches. CNN is used for image classification, followed by LSTM-based RNN for the caption generation. Bhagavatula et al. [2] investigated the viability of natural language-based abductive reasoning.

One of the obstacles on the field of text generation has been the objective evaluation of the generated text's semantics. While the spelling and grammar correctness are easy to evaluate, the meaning of the generated text can be challenging to evaluate with an automated metric [3]. Some tasks of text generation, such as image captioning or machine translation, where the (suggested) correct generation is known, can use metrics like BLEU [6] or METEOR [1], to evaluate the quality of the generated text. On the other hand, tasks where the correct generation is not specified (e.g., story [7] or poem [11] generation), the quality of the produced text is usually evaluated by human readers.

We draw inspiration for our research from the work of Luo et al. [4], who propose a novel Auto-Encoder matching model to learn utterance-level semantic dependency for generating everyday dialogue. Three neural networks are used: LSTM for encoding a reply into it's semantical representation, then a feedforward network for mapping the reply semantic into the answer semantic. Lastly a LSTM decoder is used for the sentence generation. Additionally, we found motivation for our work in the success on the field of poem generation by Potash, Romanov and Rumshisky [8], who demonstrated the generation of rap songs using LSTM, and Netzer et al. [5], who explored the usage of Word Association Norms (WAN) to generate Haiku poetry.

In the paper, we present a novel method for Haiku poetry generation, combining the approaches of generating the first verse using LSTM network, with the Auto-Encoder matching model [4] for generating the second and third verse of the Haiku poem as a form of a conversation between the verses. LSTM has proven to be capable of mimicking rhythm, common phrases and even poet's writing style [8], which we attempt to use for generating the text in a genuine Haiku style, which we describe in the continuation of the work. By using the Auto-Encoder matching model to generate the second and the third verse as an answer to the previous verse, we experiment by adding a sense of conversational bond between the verses.

2. DRUGI NASLOVI
Lorem ipsum

3. CONCLUSIONS
Lorem ipsum

4. ACKNOWLEDGMENTS

The authors acknowledge lorem ipsum.

5. REFERENCES

[1] S. Banerjee and A. Lavie. Meteor: An automatic metric for mt evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, pages 65–72, 2005.

[2] C. Bhagavatula, R. L. Bras, C. Malaviya, K. Sakaguchi, A. Holtzman, H. Rashkin, D. Downey, S. W. tau Yih, and Y. Choi. Abductive commonsense reasoning, 2020.

[3] A. Celikyilmaz, E. Clark, and J. Gao. Evaluation of text generation: A survey, 2020.

[4] L. Luo, J. Xu, J. Lin, Q. Zeng, and X. Sun. An auto-encoder matching model for learning utterance-level semantic dependency in dialogue generation, 2018.

[5] Y. Netzer, D. Gabay, Y. Goldberg, and M. Elhadad. Gaiku: Generating haiku with word associations norms. In *Proceedings of the Workshop on Computational Approaches to Linguistic Creativity*, pages 32–39, 2009.

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

[7] D. Pawade, A. Sakhapara, M. Jain, N. Jain, and K. Gada. Story scrambler-automatic text generation using word level rnn-lstm. *International Journal of Information Technology and Computer Science (IJITCS)*, 10(6):44–53, 2018.

[8] P. Potash, A. Romanov, and A. Rumshisky. Ghostwriter: Using an lstm for automatic rap lyric generation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1919–1924, 2015.

[9] S. Song, H. Huang, and T. Ruan. Abstractive text summarization using lstm-cnn based deep learning. *Multimedia Tools and Applications*, 78(1):857–875, 2019.

[10] Q. You, H. Jin, Z. Wang, C. Fang, and J. Luo. Image captioning with semantic attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[11] X. Zhang and M. Lapata. Chinese poetry generation with recurrent neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 670–680, 2014.