The 6th International Conference on Interdisciplinary research on Computer science, Psychology, and Education (ICICPE' 2022)
December 26-28, 2022. Pattaya, Thailand.

Fortune Telling via Deep Learning based Generation Model

Hyeonseok Moon[1] , Jungseob Lee[1], Jaehyung Seo[1], Sugyeong Eo[1], Chanjun Park[1,2], Jeongbae Park[1 *]

[1] Dept. of Computer Science and Engineering, Korea University
{glee889, omanma1928, seojae777, djtnrud, bcj1210, insmile}@korea.ac.kr
[2] Upstage
chanjun.park@upstage.ai

Abstract. Fortune telling refers to the work of generating an interpretation text suitable for a given fortune category (i.e. saju). Traditionally, interpreting given fortune category has been recognized as the realm of the human being, but through this study, we explore the possibility that the area of fortune telling can also be replaced by artificial intelligence. In this study, we design a deep learning-based model that generates interpretations sentences based on the types of fortune category and noun keywords to be included in the interpretation, inspired by recent studies in the field of natural language generation. We propose that, through this, it is possible to reduce the expert-level efforts required in generating fortune telling, and at the same time, to create texts with various expressions.

Keywords: Natural Language Processing, Natural Language Generation, Conditional Generation, Pretrained Language Model, Fortune Telling

1 Introduction

From a technical point of view, the fortune telling means the work of generating an interpretation text for a given fortune category (i.e. saju). Generally, each fortune category is determined by the information such as the date of birth. Then the interpretation of the corresponding category is generated by the professional human fortune tellers.

Traditionally, interpreting given fortune category has been recognized as the realm of the human being. However, we explore the possibility that the area of fortune telling can also be replaced by artificial intelligence. Through the advent of the pre-learning language models such as GPT2[1] and BART[2], requirements for the effort of human beings in natural language generation tasks, including summarization and machine translation, have largely been. This paper assumes that by applying the current natural language generation research to this, the artificial intelligence model will be able to carry out the area of solving the four weeks in which humans are currently playing a major role.

In order to utilize a deep learning model, we reconstruct the fortune telling task in a form that can be learned through the model. This means that the fortune telling is interpreted as the conditional generation task that generate interpretation text by grounding keywords that would be included in the interpretation text and fortune category. We propose that by utilizing this model, it is possible to significantly reduce the manpower required in fortune telling generation, and to generate an interpretation with a further rich expression.

2 Proposed Method

In order to perform the fortune telling task through a deep learning model, we reconstruct the fortune telling task in a form that can learn the model. In our task, language generation model is fed a fortune category and keywords to be included in the interpretation. Then the model is trained to generate the corresponding fortune telling interpretation text. This task is inspired by the common sense text generation task (CommoGgen)[3], where several noun and verb keywords are fed into the model as input and model is argued to generate fluent sentences that include all of the given keywords.

To this end, we segment the fortune telling interpretation text into sentences and extract keywords included in each sentence. To extract keywords in sentences, we utilize mecab-ko[1], a Korean morpheme analysis framework. After that, a single input sequence is constructed by concatenating both the noun keyword and the fortune category, and based on this, the model is trained to generate the related fortune telling interpretation text.

3 Experimental Setting

The data used in this study were provided by the FLES, which operates the fortune telling contents, for research purpose. The data consisted of a total of 36.862 sentences and a total of 171 fortune categories, and we used compose train, valid, test split by applying 8:1:1 ratio. We used BART and GPT2 as pre-trained language models, and all learning was conducted using a single A6000 GPU.

4 Main Results

Table 1. Experimental results

	BLEU score	BERT score
GPT2	16.629	57.509
BART	53.914	76.380

[1] https://bitbucket.org/eunjeon/mecab-ko-dic/src/master/

We confirmed the generation result using the given test dataset. To evaluate the performance of the model, we generated 20 interpretation sentences through a beam search decoding strategy. Then, the performance was evaluated by comparing the average score of the bleu score and the BERT score between these sentences and the reference sentence. As can be seen from the experimental results, it was difficult to see that GPT2 generates sentences that are highly related to the correct answer, but it can be confirmed that the model using BART can generate sentences with high correlation. This shows that our proposal can create artificial intelligence models that replace human effort.

5 Conclusion

We proposed a deep learning-based fortune telling model the generate interpretation text through the keywords and fortune category. By combining the technologies of CommonGen research and generative control research, which are currently being studied, we replaced the an artificial intelligence system that generates the corresponding keynote answer based on the keynote type and noun keyword. As a future study, we plan to conduct research to apply various decoding strategies.

5 Acknowledgement

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. 2020-0-00368, A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques) and the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2022-2018-0-01405) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

References

1. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.
2. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2020, July). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 7871-7880).
3. Lin, B. Y., Zhou, W., Shen, M., Zhou, P., Bhagavatula, C., Choi, Y., & Ren, X. (2019). CommonGen: A constrained text generation challenge for generative commonsense reasoning. arXiv preprint arXiv:1911.03705.