

Enhancing Machine Translation with Quality Estimation and Reinforcement Learning

Zijian Győző Yang, László János Laki

Hungarian Research Centre for Linguistics
{yang.zijian.gyozo,laki.laszlo}@nytud.hu

Abstract

"Machine translation is the process of automatically translating text from one language to another. In recent years, advances in artificial intelligence and deep learning have led to significant improvements in the quality of machine translation. Reinforcement learning is a type of machine learning that focuses on training models to make a sequence of decisions based on rewards. In the context of machine translation, reinforcement learning has been used to fine-tune machine translation models and incorporate human feedback into the training process. Quality estimation, on the other hand, is the task of predicting the quality of machine-translated text without having access to a reference translation. This is a crucial step in the machine translation pipeline as it allows the system to determine which translation is the most accurate and choose the best one to return to the user. By combining reinforcement learning and quality estimation, machine translation systems can continuously improve and provide high-quality translations."

The first paragraph is generated by ChatGPT¹, which correctly summarizes the current research. ChatGPT is fine-tuned from a model in the GPT-3.5 series [1], which is one of the hottest topic in the recent months. The enormous success is attributed to the reinforcement learning method, which integrates human feedback into the language modeling process.

In our research, we have successfully applied a neural quality estimation model and the reinforcement learning method into the machine translation training process to improve its quality.

¹<https://chat.openai.com>

Based on research by OpenAI [4], the steps for using reinforcement learning in fine-tuning language models are as follows:

1. *Fine-tuning language model with supervised learning:* In our experiment, we used a mT5 small model² [3], that fine-tuned for English-Hungarian translation task (supervised fine-tuned model - SFT).
2. *Collect human feedback and training a reward model:* For this task, we trained a quality estimation model (as reward model) for English-Hungarian translation. We fine-tuned a XLM-RoBERTa base model [2] for quality estimation (XLM-QE) task on the HuQ [6] corpus.
3. *Fine-tuning language model with reward model and reinforcement learning method:* We further fine-tuned the SFT mT5 model with XLM-QE and reinforcement learning method (RL mT5). For fine-tuning we used the CarperAI implementation³ and for training and test set we used the official subcorpus of Hungarian [5] from Shared Task of WMT 2009⁴.

In Table 1, you can see that our new quality estimation model, the XLM-QE significantly (>20% correlation) outperformed the baseline model (published in the research of Yang et. al [6]) and achieved state-of-the-art results in English-Hungarian quality estimation task.

Table 1. Results of the quality estimation task

	Correlation	MAE	RMSE
baseline	0.6100	0.7459	0.9775
XLM-QE	0.8382	0.5785	0.8184

In Table 2, you can see the results of SFT and RL mT5 machine translation models. Our RL mT5 model could significantly outperformed the SFT mT5 model (>5 BLEU score). One of the main problem of the original SFT mT5 model is that the generated outputs are longer than the source text and contain incorrect repeated phrases (as shown in Table 3). This results in high recall but low precision. By utilizing a human-based reward model and reinforcement learning method, we were able to correct these errors (as evidenced in Table 3).

²<https://huggingface.co/NYTK/translation-mt5-small-128-en-hu>

³<https://github.com/CarperAI/trlx>

⁴<https://www.statmt.org/wmt09/translation-task.html>

Table 2. Results of the RL mT5 model

	BLEU	chrF-3	chrF-6
SFT mT5	7.34	45.60	38.62
RL mT5	12.91	47.02	40.39

Table 3. A sample of the correction of RL mT5 model

Source:	A signal for Asian trading
Reference:	Jelzés az ázsiai kereskedésnek
SFT mT5:	Egy jel az ázsiai kereskedésre vonatkozóan: jel ázsiai kereskedésre
RL mT5:	Egy jel az ázsiai kereskedéshez

References

- [1] T. BROWN, B. MANN, N. RYDER, M. SUBBIAH, J. D. KAPLAN, P. DHARIWAL, A. NEELAKANTAN, P. SHYAM, G. SASTRY, A. ASKELL, S. AGARWAL, A. HERBERT-VOSS, G. KRUEGER, T. HENIGHAN, R. CHILD, A. RAMESH, D. ZIEGLER, J. WU, C. WINTER, C. HESSE, M. CHEN, E. SIGLER, M. LITWIN, S. GRAY, B. CHESS, J. CLARK, C. BERNER, S. McCANDLISH, A. RADFORD, I. SUTSKEVER, D. AMODEI: *Language Models are Few-Shot Learners*, in: Advances in Neural Information Processing Systems, ed. by H. LAROCHELLE, M. RANZATO, R. HADSELL, M. BALCAN, H. LIN, vol. 33, Curran Associates, Inc., 2020, pp. 1877–1901.
- [2] A. CONNEAU, K. KHANDELWAL, N. GOYAL, V. CHAUDHARY, G. WENZEK, F. GUZMÁN, E. GRAVE, M. OTT, L. ZETTLEMOYER, V. STOYANOV: *Unsupervised Cross-lingual Representation Learning at Scale*, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online: Association for Computational Linguistics, July 2020, pp. 8440–8451, doi: [10.18653/v1/2020.acl-main.747](https://doi.org/10.18653/v1/2020.acl-main.747).
- [3] L. J. LAKI, Z. G. YANG: *Neural machine translation for Hungarian*, Acta Linguistica Academica 69.4 (2022), pp. 501–520, doi: <https://doi.org/10.1556/2062.2022.00576>.
- [4] L. OUYANG, J. WU, X. JIANG, D. ALMEIDA, C. WAINWRIGHT, P. MISHKIN, C. ZHANG, S. AGARWAL, K. SLAMA, A. GRAY, J. SCHULMAN, J. HILTON, F. KELTON, L. MILLER, M. SIMENS, A. ASKELL, P. WELINDER, P. CHRISTIANO, J. LEIKE, R. LOWE: *Training language models to follow instructions with human feedback*, in: Advances in Neural Information Processing Systems, ed. by A. H. OH, A. AGARWAL, D. BELGRAVE, K. CHO, 2022.
- [5] D. VARGA, P. HALACSY, A. KORNAI, V. NAGY, L. NEMETH, V. TRON: *Parallel corpora for medium density languages*, in: Recent Advances in Natural Language Processing IV. Selected papers from RANLP-05, ed. by N. NICOLOV, K. BONTCHEVA, G. ANGELOVA, R. MITKOV, Amsterdam: Benjamins, 2007, pp. 247–258.
- [6] Z. G. YANG, J. L. LAKI, B. SIKLÓSI: *HuQ: An English-Hungarian Corpus for Quality Estimation*, in: Proceedings of the LREC 2016 Workshop - Translation Evaluation: From Fragmented Tools and Data Sets to an Integrated Ecosystem (Portorož, Slovenia, May 24, 2016), 2016.