

СОВМЕСТНОЕ ОБУЧЕНИЕ ПРЯМОЙ И ОБРАТНОЙ МОДЕЛЕЙ МАШИННОГО ПЕРЕВОДА

Скачков Николай Андреевич

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: nikolaj-skachkov@yandex.ru

Научный руководитель — Воронцов Константин Вячеславович

Задача машинного перевода является одной из важнейших задач анализа текстов естественного языка. Перевод текстов является достаточно трудоёмким процессом, затрачивающим существенное количество человеческого труда и времени.

Современные модели перевода использует нейросетевые модели, обученные на параллельных предложениях. Эти модели на основе большого количества параллельных текстов выучивают межязыковые закономерности и используют их при переводе. В основе большинства современных моделей лежит архитектура Transformer [1].

Хорошим критерием качества перевода является то, что обратный перевод через обратную модель, например русско-английскую для англо-русской задачи, должен совпадать с исходным текстом. Это означает, что при переводе туда-обратно не было добавлено или потеряно смысла в сравнении с исходным предложением. В работе [2] было предложен метод обучения с использованием обратной модели, а также дополнительных языковых моделей исходного и целевого языков. Сложность такого подхода заключается в том, что предсказания модели являются текстами на целевом языке, то есть дискретны. Следовательно, необходимо решить задачу как пробросить через дискретные переменные градиенты по функции потерь от обратной модели. Общая схема модели в исходной статье достаточно сложная и трудновоспроизводимая.

В наших экспериментах удалось избежать использования языковых моделей, а также сделать обучение существенно более стабильным. Для этого были выведены формулы для проброса градиентов функции потерь обратной модели с помощью метода REINFORCE [3]. Также вместо обучения с нуля такой модели использовалась предобученная стандартным способом модель перевода. Такое решение позволило решить проблему нестабильности метода REINFORCE и позволило эффективно дообучить модель. Подтверждением этого является понижение средней нормы градиентов при дообучении более чем в 10 раз.

При этом, нижняя оценка функции потерь при использовании обратной модели выглядит следующим образом:

$$\mathcal{L}(\Theta, \text{src}) = \sum_{\text{dst}} P_{\Theta}(\text{dst}|\text{src}) \log P_{\Phi}(\text{src}|\text{dst}),$$

где Φ — обратная модель, Θ — прямая модель, а src, dst — это исходный текст и перевод, причём суммирование ведётся по всевозможным переводам. Поскольку предсказания обратной модели являются константой относительно параметров модели, то данную функцию потерь можно воспринимать как взвешенную сумму, в которой будут повышаться вероятности тех переводов, у которых обратный перевод совпадает с исходным текстом.

Для получения градиентов по этой функции используется метод REINFORCE.

В экспериментах удалось добиться прироста по основной метрике машинного перевода BLEU [4] на англо-финском и русско-казахском переводах относительно стандартного обучения. Прирост составляет порядка 0.5 BLEU в обоих экспериментах, что является значимым приростом. При этом модель предложенную в [2] не удалось обучить из-за нестабильности и большой дисперсии градиентов. Таким образом, предложенный подход позволяет улучшить качество перевода, за счёт достаточно упрощённой процедуры дообучения.

Литература

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I. // Attention Is All You Need. NIPS, 2017
2. Di He¹, Yingce Xia², Tao Qin³, Liwei Wang¹, Nenghai Yu², Tie-Yan Liu³, Wei-Ying Ma. Dual Learning for Machine Translation
3. Ivanov S., D'yakonov A. Modern Deep Reinforcement Learning Algorithms
4. Papineni, Kishore Roukos, Salim Ward, Todd Zhu, Wei-Jing. Bleu: a Method for Automatic Evaluation of Machine Translation, // Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, 2002, P. 311-318a