

ОБЗОР-РЕЦЕНЗИЯ НА СТАТЬЮ

"What Can Transformers Learn In-Context? A Case Study of Simple Function Classes"

Зинов Александр

Высшая Школа Экономики

aazinov@edu.hse.ru

7 октября 2022 г.

Как все было?

OpenAI сделали GPT-3 и понеслось. Оказалось, что одна модель может в зависимости от текстового контекста решать очень разные задачи:

- Tom B. Brown и др. *Language Models are Few-Shot Learners*. 2020. arXiv: 2005.14165 [cs.CL]

Таким образом можно аппроксимировать байесовские модели:

- Samuel Müller и др. *Transformers Can Do Bayesian Inference*. 2021. DOI: 10.48550/ARXIV.2112.10510. URL: <https://arxiv.org/abs/2112.10510>

Почему бы и не другие модели?

Кто авторы?

- Dimitris Tsipras - post-doctoral researcher, Stanford

Под руководством:

- Percy Liang
- Greg Valiant

- Shivam Garg - PhD, Stanford

Под руководством:

- Greg Valiant

Публикуют статьи про разное, до этого про трансформеры и meta learning не писали.

В чем суть?

- Продолжаем и формализуем до готовности идею о том, что трансформеры можно научить учиться.
- Для предсказания значения функции из заранее выученного класса не нужно градиентного спуска и большой обучающей выборки.
- Таким образом хотим лучше осознать наблюдение, что текстовые трансформеры умеют в in-context learning.

Плюсы статьи

- Формализует common knowledge на горячую тему.
- Статья сопровождается хорошо структурированным кодом.
- Грамотно демонстрирует, что для небольших моделей рассматриваемый подход позволяет получить SOTA.

Минусы статьи

- Практического применения метод, предложенный в статье, скорее всего не имеет (к этому еще вернемся).
- Не хватает чётко сформулированных выводов. Авторы ставят своей целью лучше понять умение текстовых трансформеров решать задачи, на которые они не обучались, но по итогу никак не обсуждают этот вопрос в тексте, ограничиваясь связанными примерами для других моделей.

А как же PTuning?

На практике in-context learning проигрывает PTuning или Finetuning по качеству и особенно производительности:

- Haokun Liu и др. *Few-Shot Parameter-Efficient Fine-Tuning is Better and Cheaper than In-Context Learning*. 2022. DOI: [10.48550/ARXIV.2205.05638](https://doi.org/10.48550/ARXIV.2205.05638). URL: <https://arxiv.org/abs/2205.05638>

И все таки, почему это работает?

Вопрос умения текстовых трансформеров обучаться на новые задачи при инференсе на данный момент широко исследуется. Вот пример косвенно конкурирующей статьи, в которой данный феномен исследуется более непосредственно:

- Sang Michael Xie и др. *An Explanation of In-context Learning as Implicit Bayesian Inference*. 2021. DOI: [10.48550/ARXIV.2111.02080](https://arxiv.org/abs/2111.02080). URL: <https://arxiv.org/abs/2111.02080>

Статья новая, поэтому цитирование пока только одно. Авторы повторяют трюк для RNN и даже формулируют верхнюю оценку на ошибку in-context learner'a:

- [Surbhi Goel и др.](#) *Recurrent Convolutional Neural Networks Learn Succinct Learning Algorithms*. 2022. DOI: [10.48550/ARXIV.2209.00735](https://arxiv.org/abs/2209.00735). URL: <https://arxiv.org/abs/2209.00735>

Что дальше предлагают авторы?

Авторы предполагают, что структура трансформера позволяет понять, как именно обученная модель справляется с обученным классом задач. Возникает интересная идея поискать таким образом новые алгоритмы для известных задач.

Что дальше предлагаю я?

Как известно, для традиционных текстовых трансформеров важен порядок передачи токенов в модель. Было бы интересно исследовать вопрос влияния порядка сэмплов на качество модели, обученной предложенным методом.

- Yao Lu и др. *Fantastically Ordered Prompts and Where to Find Them: Overcoming Few-Shot Prompt Order Sensitivity*. 2021. DOI: 10.48550/ARXIV.2104.08786. URL: <https://arxiv.org/abs/2104.08786>