

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

Shivam Garg, Dimitris Tsipras, Percy Liang, Gregory Valiant

Выполнил Макаров Г.Н.
БПМИ 193

In-context learning

maison → house, chat → cat, chien → dog
prompt completion

In-context learning / обучение по контексту – модель без fine-tuning способна решать новую задачу, получая на вход лишь небольшую подсказку.

Пример выше, для сложных моделей: GPT-3 по силам даже сделать перевод с французского на английский

Линейные функции из распределения $D_{\mathcal{F}}$

x_i, x_{query} из распределения $D_{\mathcal{X}}$

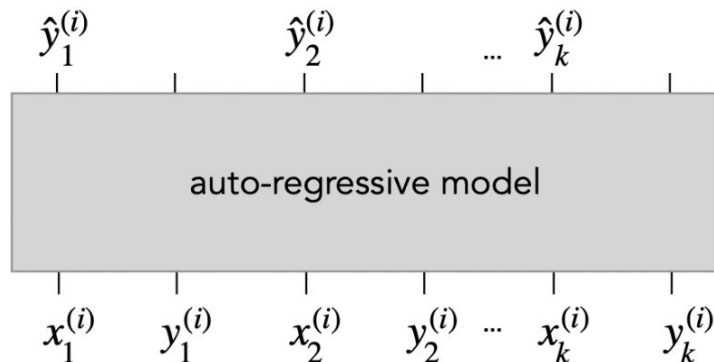
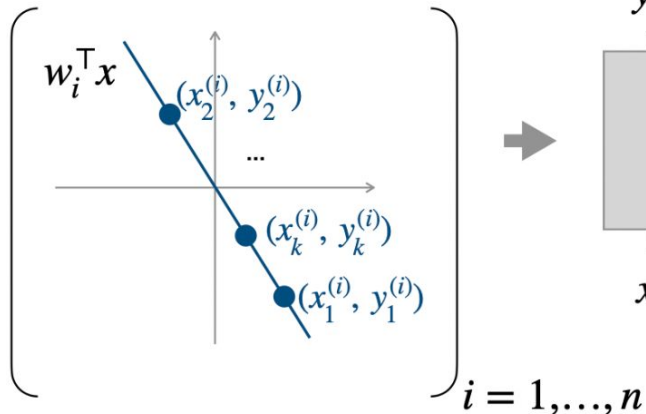
$(x_1, f(x_1), \dots, x_k, f(x_k), x_{\text{query}})$ - последовательность подсказок

$$\mathbb{E}_P [\ell (M (P) , f (x_{\text{query}}))] \leq \epsilon,$$

Линейные функции

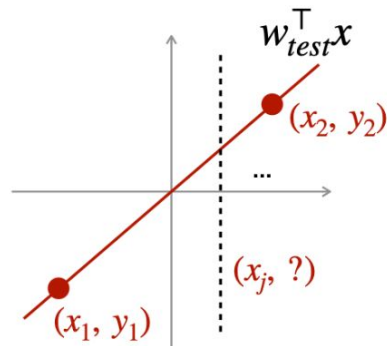
Training data

$$w_1, \dots, w_n \stackrel{i.i.d.}{\sim} N(0, I_d)$$



Inference

$$w_{test} \sim N(0, I_d)$$

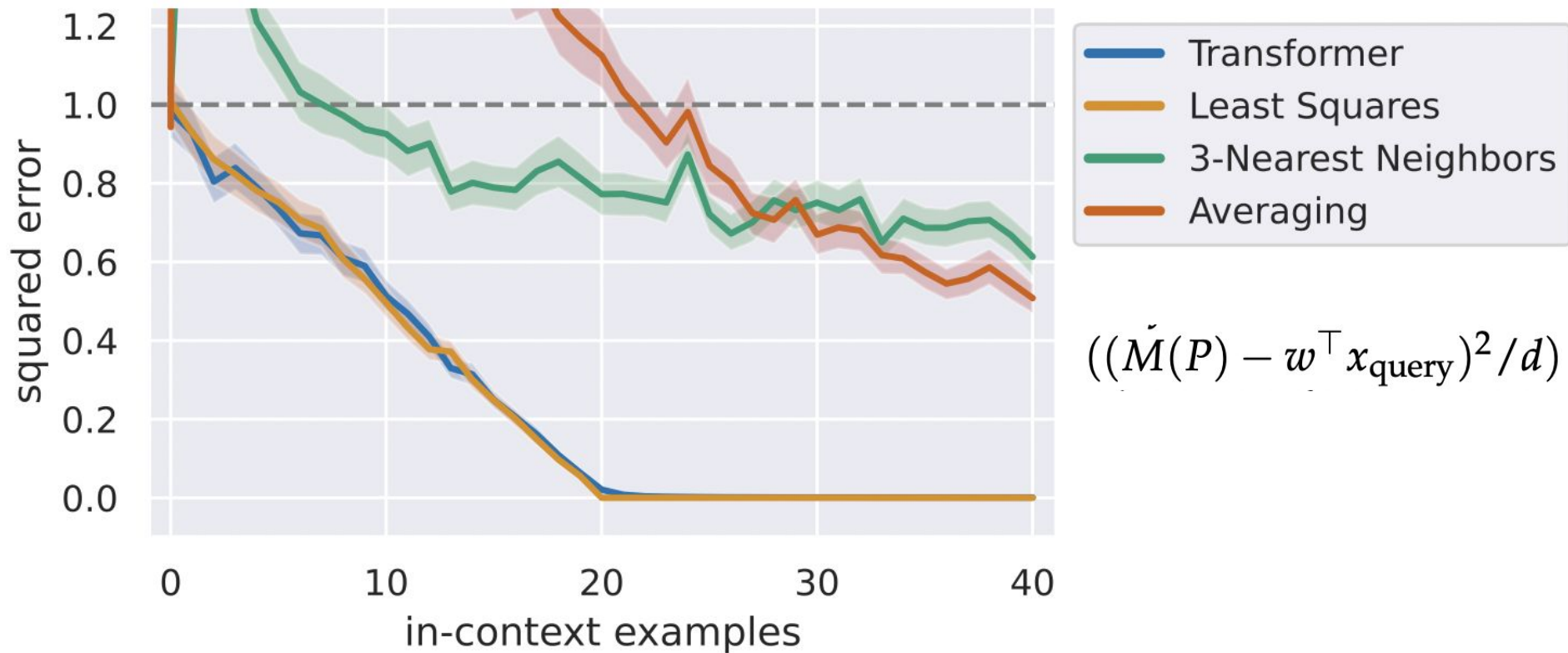


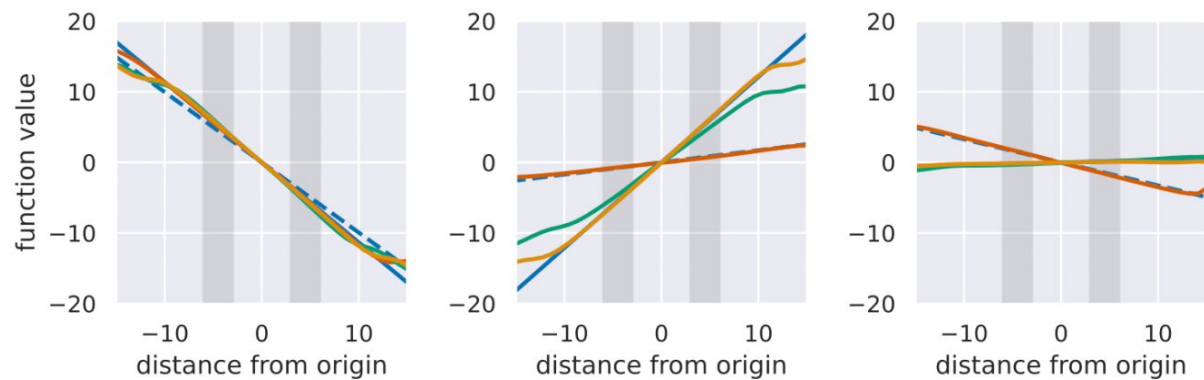
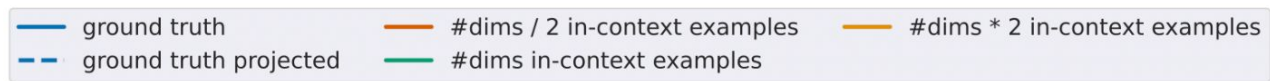
Пример того, как мы обучаем модель определять линейную зависимость

Квадратичная формула обучения

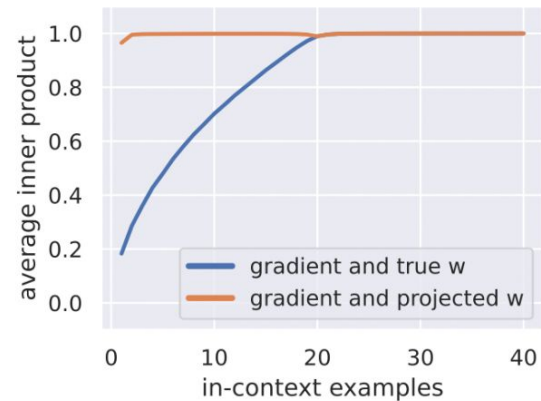
$$\min_{\theta} \mathbb{E}_P \left[\frac{1}{k+1} \sum_{i=0}^k \ell \left(M_{\theta} \left(P^i \right), f \left(x_{i+1} \right) \right) \right]$$

Квадратичная ошибка модели Transformer на линейных функциях



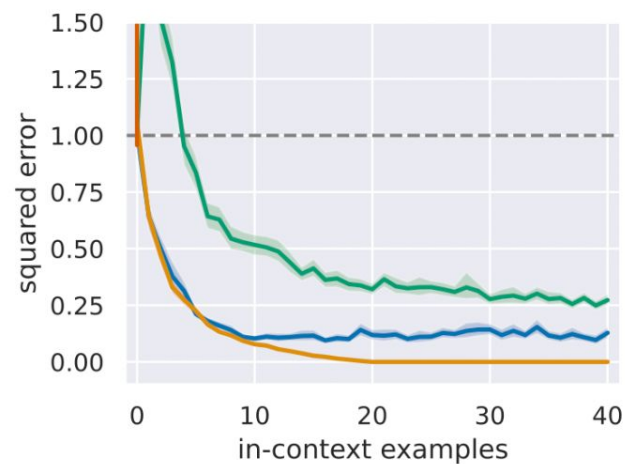


(a) function visualizations

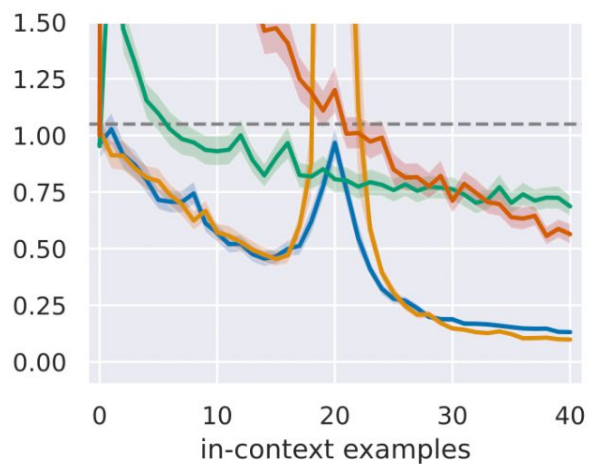


(b) gradients

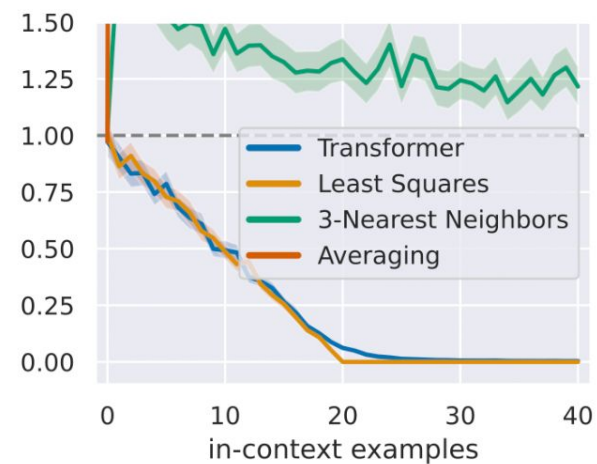
Сдвиги



(a) Skewed covariance

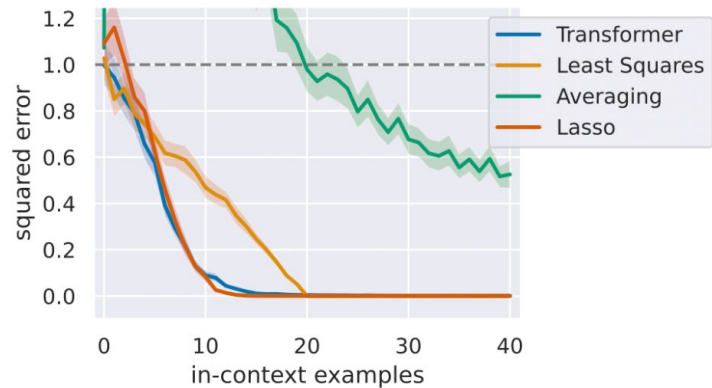


(b) Noisy linear regression

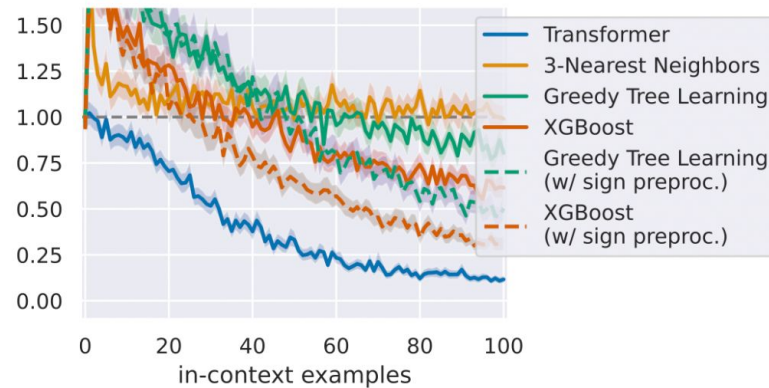


(c) Different orthants

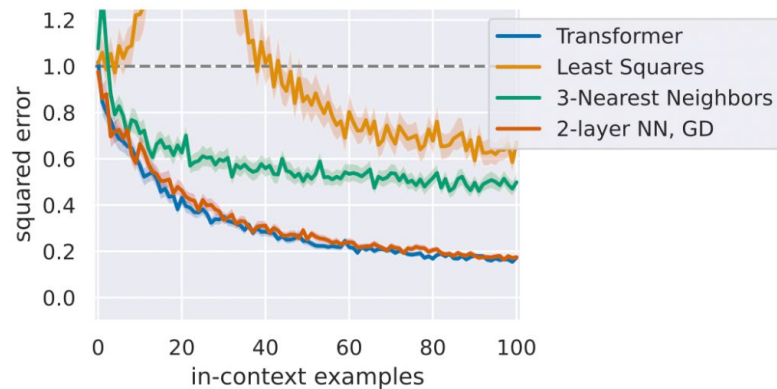
Более сложные функции



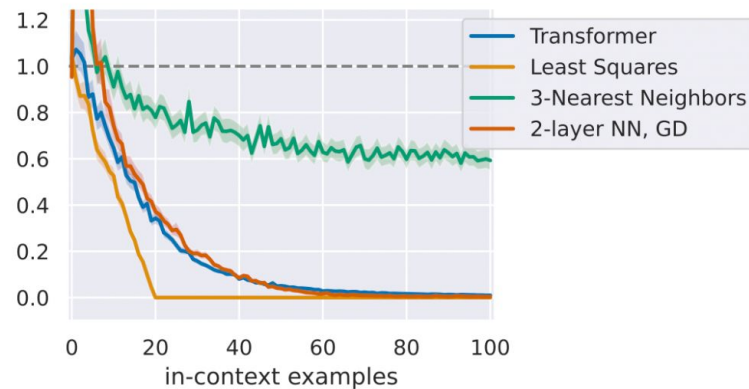
(a) Sparse linear functions



(b) Decision trees

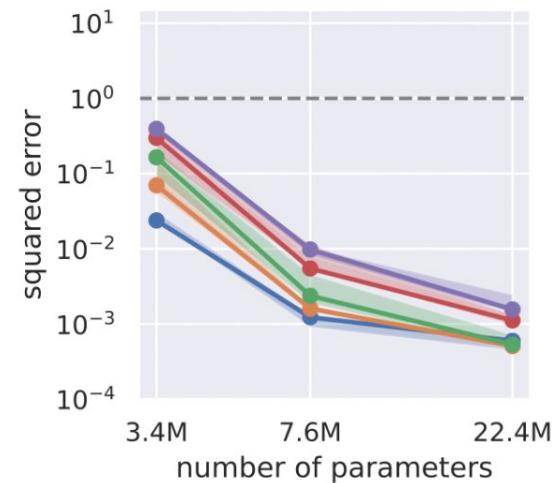


(c) 2-layer NN

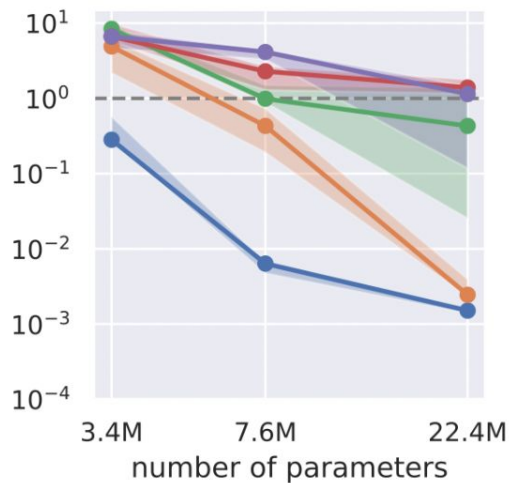


(d) 2-layer NN, eval on linear functions

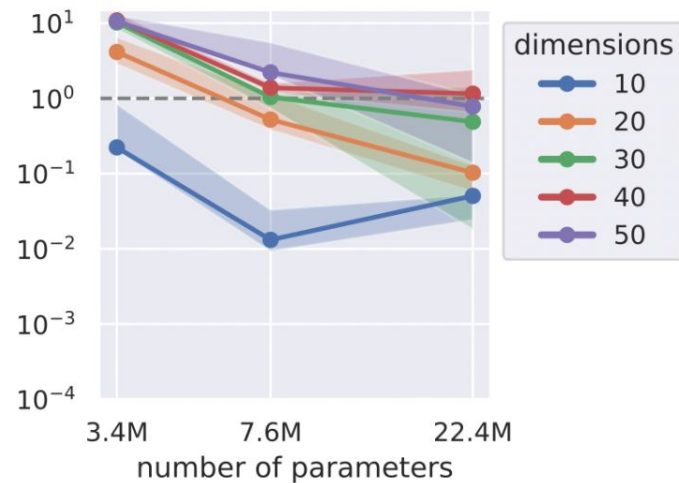
In/Out distribution подсказок



(a) Standard



(b) Different orthants



(c) Skewed covariance

Спасибо за внимание!

Обзор научного контекста

Выполнил Карлов В.А.
БПМИ193

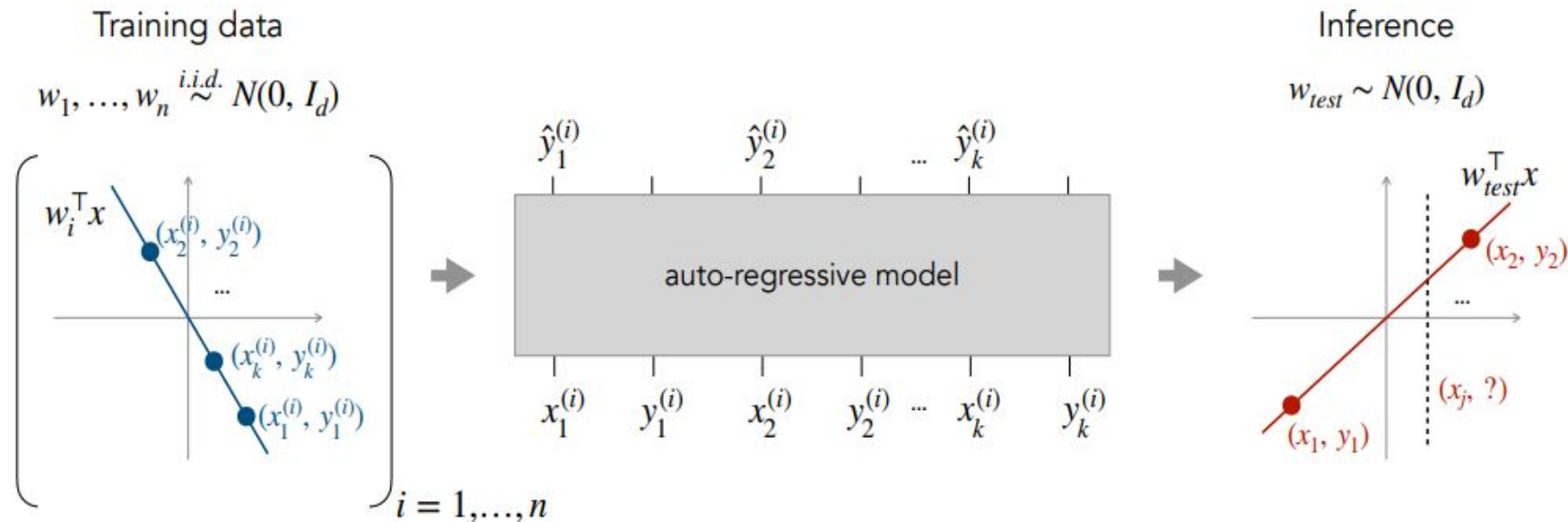
Суть работы

$\underbrace{\text{maison} \rightarrow \text{house}, \text{chat} \rightarrow \text{cat}, \text{chien} \rightarrow}_{\text{prompt}} \underbrace{\text{dog}}_{\text{completion}} .$

In-context learning / обучение по контексту – модель без fine-tuning'а способна решать новую задачу, получая на вход лишь небольшую подсказку.

В примере выше – перевод с французского на английский.

Суть работы



Пример того, как это выглядит для линейных функций

Про публикацию

- Опубликовано 1 августа 2022 года
- Только у одного из исследователей есть другая работа, связанная с in-context learning, поэтому статья является скорее случайно находкой, нежели продолжением серии изысканий

Про публикацию

- 4 автора из Стенфорда (2 доцента и 2 аспиранта, находящихся под их руководством)



Shivam Garg
40 цитирований



Dimitris Tsipras
4950 цитирований



Percy Liang
9875 цитирований



Gregory Valiant
674 цитирований

Related work – Litmap

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

Garg, 2022 – ArXiv

Top 20 related articles

Save...

PaLM: Scaling Language Modeling with Pathways

Chowdhery, 2022 – ArXiv

Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Min, 2022 – Conference on Empirical Methods in Natural Language Processing

Emergent Abilities of Large Language Models

Wei, 2022 – ArXiv

What learning algorithm is in-context learning? Investigations with linear models

Akyürek, 2022 – ArXiv

In-context Learning and Induction Heads

Olsson, 2022 – ArXiv

Transformers learn in-context by gradient descent

Oswald, 2022 – ArXiv

OPT: Open Pre-trained Transformer Language Models

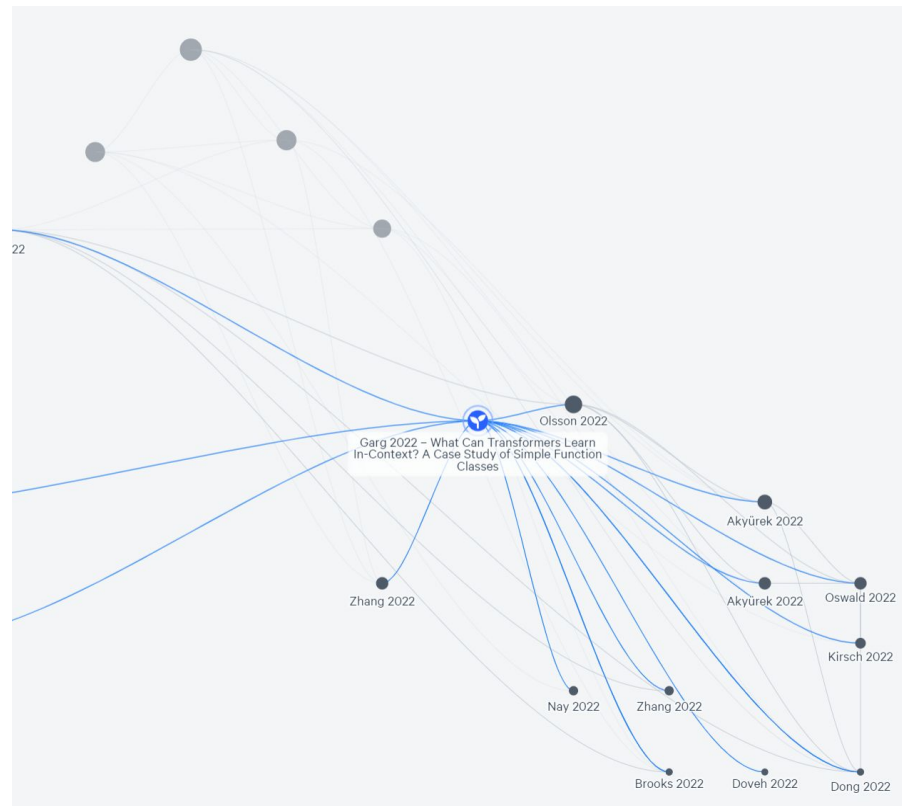
Zhang, 2022 – ArXiv

Training Compute-Optimal Large Language Models

Hoffmann, 2022 – ArXiv

A Survey for In-context Learning

Dong, 2022 – ArXiv



Related work

Можно выделить 4 связанных с работой области:

1. In-context learning
2. Transformers
3. Meta learning (learning to learn)
 - a. учим “meta-learner” обновлять параметры при обучении на разные последующие задачи
 - b. учим значения параметров, из которых можно решать большое количество разных типов задач
4. Data-driven algorithm design

Related work / In-context Learning

1. “Language Models are Few-Shot Learners” – Brown et al. 2020
 - a. Известная и объемная работа
 - b. Одна из ключевых для in-context

2. “An explanation of in-context learning as implicit bayesian inference” – Xie et al. 2022
 - a. Все авторы из Стенфорда, один общий доцент
 - b. Рассматривает контекстное обучение как скрытый байесовский вывод
 - c. Предлагают фреймворк для байесовского вывода, который объясняет, как работает контекстное обучение

Related work / Transformers

1. “Attention is all you need” – Vaswani et al. 2017 и другие
2. “Transformers can do bayesian inference” – Müller et al. 2021
 - a. вводят “Prior-data fitted transformer”
 - b. Обучаем аппроксимировать байесовский вывод с гауссовскими процессами и байесовскими нейронными сетями в качестве априорных распределений
 - c. полученные знания используются для решения разных задач, например, классификации на табличных данных или few-shot классификации картинок
 - d. по сути очень похоже на TabPFN

Цитирования и продолжения

Цитирования:

- 7 на Google Scholar
- 17 на Semantic Scholar

Развития работы:

- “General-Purpose In-Context Learning by Meta-Learning Transformers”,
8 декабря 2022 года
- “Why Can GPT Learn In-Context? Language Models Secretly Perform Gradient Descent as Meta-Optimizers”,
20 декабря 2022 года

Сильные стороны

1. Наличие исходного кода в открытом доступе
2. Основательность, с которой авторы подошли к исследованию и проведению экспериментов
 - a. зависимость от данных, на которых обучается модель
 - b. устойчивость контекстно обученных трансформеров к сдвигам распределений подсказок
 - c. факторы, влияющие на контекстное обучение (размерность задачи и размер модели)
3. Inductive bias трансформеров, как архитектуры (какие классы функций им сложнее / проще выучивать по контексту)
4. Обширная формализация, почти нет лишних обозначений

Слабые стороны

1. Слабо структурированный демонстрационный irunb-ноутбук с небольшим количеством комментариев
2. Мало примеров про контекстное обучение, из-за чего в работу тяжело быстро вникнуть
3. Есть повторения в тексте