

# Рецензия

Статья исследует популярный метод сжатия / ускорения нейронных сетей – дистилляцию, опираясь на обобщающую способность и согласованность (насколько точно совпадают выход студента и учителя). Авторы провели большое количество экспериментов и выдвинули на основе этого несколько гипотез. Ключевым открытием является отсутствие корреляции между обобщающей способностью (которую измеряли с помощью ассигасы) и согласованностью (fidelity).

## Сильные стороны

1. Другие работы о дистилляции знаний предлагали свои подходы, исследовали, какой должна быть сеть-учитель и так далее. Однако какого-то общего исследования данного метода (работает ли он как ожидается) до этой статьи не было.
2. Статья актуальна, ведь задача сжатия нейронных сетей активно исследуется в наши дни. Это позволяет ускорить работу нейронной сети и снизить ресурсоемкость модели.
3. Авторы провели обширный эмпирический анализ, эксперименты были проведены на разных данных (включая разные типы данных: картинки, текст).

## Слабые стороны

1. Нет предложений как можно улучшить процедуру обучения модели студента, несмотря на интересные гипотезы в результате эксперимента.
2. Исследования проведены только с классическим видом дистилляции знаний, однако, возможно, другие варианты вели бы себя по другому.

## Качество статьи

Статья написана на доступном языке, текст хорошо структурирован. Все сложные фрагменты сопровождаются комментариями и пояснениями.

## Воспроизводимость

Авторы предоставляют все необходимые детали для воспроизведения экспериментов. Кроме того, имеется хорошо организованный репозиторий с исходным кодом всех экспериментов на Github.

## Оценка статьи

Оценка: 7 (A good submission; an accept).

Уверенность: 4 (You are confident in your assessment, but not absolutely certain).

