

Название статьи (авторы статьи): Does Knowledge Distillation Really Work?

Автор исследования: Алекберов Артем

При проведении исследования, мы предлагаем ответить вам на следующие вопросы:

*1. Когда написана работа? Опубликована ли она на какой-то конференции? Если да, то как она была представлена на конференции (орал/спотлайт/постер)? Есть ли у нее какая-то история публикации (может быть ее взяли не сразу)?*

Работа была принята на NeurIPS 2021. Оценки на openreview были по большей части положительные, рецензенты отмечали хорошую экспериментальную базу, а также важность итоговых выводов статьи для понимания работы KD.

*2. Кто авторы статьи? Где и над чем они работают? Как другие проекты авторов связаны с этой статьей? Основываясь на тексте статьи и истории публикаций авторов, подумайте как авторы пришли к идее статьи -- может быть это прямое улучшение их предыдущей работы, а может это больше выглядит как случайная находка.*

Большая часть авторов из Нью-Йоркского Университета, из научной группы Andrew Gordon Wilson. Авторы в основном занимаются вероятностными методами в глубинном обучении, байесовскими методами, проблемой генерализации и другими теоретическими вопросами. До этого никто из авторов не углублялся в вопрос KD, но позиция теоретической обоснованности уже известных алгоритмов четко обозначается на персональных страницах авторов. Также можно обратить внимание на то, что двое из авторов - Павел Измайлов и Полина Кириченко - бывшие резиденты Байесовской группы.

*3. Какие из статей в списке ссылок оказали наибольшее влияние на данную работу? Можно ли выделить какие-то 1-3 статьи, которые можно назвать базовыми для этой работы? Опишите в чем связь с этими работами (без математики, просто суть).*

Авторы не сравниваются с какой-то конкретной статьей, скорее наоборот взяли миссию происследовать моменты, ранее пропущенные и не обговоренные в статьях про дистилляцию. Упоминаются статьи Fast, Accurate, and Simple Models for Tabular Data via Augmented Distillation за улучшение дистилляции с помощью аугментаций в табличных данных, Ensemble Distribution Distillation за успешный подход дистилляции ансамблей и Knowledge distillation: A good teacher is patient and consistent за исследование техник для улучшения качества целевой задачи с помощью KD. Авторы статьи же рассматривают скорее перпендикулярную постановку - почему студент не может сравниться с учителем даже на тех же данных, на которых они оба учились, коррелирует ли это с целевой задачей и много других теоретических вопросов.

*4. Кто цитирует данную статью? Есть ли у этой работы прямые продолжения, которые стоит прочесть тем, кто заинтересовался этой работой?*

На данный момент прямых последователей нет. Статью цитируют в surveys и других статьях в качестве одной из важных статей для понимания дистилляции, но, например, авторы *Adaptive Distillation: Aggregating Knowledge from Multiple Paths for Efficient Distillation* используют метрику для оценки "fidelity" из статьи, что в общем-то и было одной из целей статьи - донести важность данной метрики для оценки качества дистилляции.

*5. Попробуйте на основе результатов статьи предложить исследование, не проведенное к текущему моменту.*

Известно, что использование других обджективов (MSE/MAE на логиты) или intermediate layers distillation значительно улучшают качество дистилляции. Отсюда хотелось бы понять распространяются ли выводы статьи на использование данных методов?

Также хочется дальше поисследовать, не является ли важность аугментаций результатом переобучения модели? Авторы говорят об улучшении качества на тесте, а также упоминают, что качество на трейне стало хуже, что косвенно подтверждает осмысленность данного исследования.

Также авторы говорят о важности инициализации близкой к оптимальному учителю. Отсюда появилась мысль попробовать воссоздать инициализацию учителя в более маленьком студенте через некоторую проекцию, а затем попробовать дистилляцию.

*6. Попробуйте найти применение статье в индустриальных приложениях.*

1. Стоит помнить, что использование аугментаций или другое расширение датасета являются важной составляющей для достижения хорошего результата дистилляции.
2. Стоит учитывать и использовать введенные авторами метрики для учета качества того, насколько студент смог повторить учителя.