

Sharpness-aware Minimization for Efficiently Improving Generalization

1. Содержание и вклад

В статье описан метод оптимизации, который учитывает особенности поверхности, которую образует функция потерь, и связь между обобщающей способностью и остротой экстремума.

Функции потерь у современных нейросетей образуют сложные невыпуклые поверхности, из-за чего выбор метода оптимизации — важное архитектурное решение. Также есть эмпирические исследования о том, что модель, сошедшаяся к более пологому минимуму, имеет лучшую обобщающую способность. Метод, который учитывает эти особенности, может повысить эффективность обучения и улучшить обобщающую способность моделей.

2. Сильные стороны

- Проведено множество экспериментов на разных наборах данных с разными архитектурами и методами регуляризации. Также рассмотрены разные условия: обучение модели с нуля, дообучение, обучение на шумных данных.
- Эксперименты показали действительно хорошие результаты. Практически во всех экспериментах удалось улучшить качество предсказаний относительно результата, полученного без использования SAM.
- Статья написана четко и понятно, ее легко читать. Сам текст хорошо структурирован.
- Метод легко воспроизводится по материалам из статьи, так как в тексте приведены все подробности. Помимо этого в открытом доступе есть авторская реализация. Подробнее о воспроизводимости в соответствующей [секции](#).

3. Слабые стороны

- Нет экспериментальных сравнений с близкими работами [1, 2, 3] и прямыми конкурентами [1].

- Схожий метод уже был предложен под названием extragradient method в 1976 году. Этот метод неоднократно появлялся в работах по оптимизации для глубинного обучения [[1](#), [2](#)] в последние годы, хоть и для других целей. В любом случае, из-за близости методов это критически важное сравнение.
- Недостаточно изучено поведение метода в конкретных условиях. Проведенные эксперименты фокусируются только на улучшении метрик. Само поведение метода не очень хорошо исследовано.
- У метода нет теоретического обоснования, что может быть важно для метода оптимизации. Шаги методы были выведены из оценки сверху в условиях [PAC learning](#). В условиях PAC learning делается много довольно жестких предположений о задаче, поэтому на практике PAC оценки не держатся. Таким образом, метод не обоснован.

4. Качество текста

Статья написана довольно доходчиво. Отрывков, требующих доработки, нет, поскольку все написано четко и однозначно.

5. Воспроизводимость

Статья написана достаточно подробно для того, чтобы ее результаты можно было повторить.

В статье подробно расписана экспериментальная установка (использованное железо, процесс подбора гиперпараметров, замер качества модели). Также приведены рекомендованные значения гиперпараметров, а сам код выложен в [открытый доступ](#).

Помимо официального репозитория статьи существуют неофициальные реализации [[1](#), [2](#)], в которых попытались провести эксперименты с предложенным методом. В наиболее популярном репозитории удалось воспроизвести результат статьи.

6. Дополнительные комментарии

Вероятно, следует провести больше сравнений со схожими работами, часть из которых приведена в [секции о слабых сторонах работы](#). Также может быть полезно более подробно изучить поведение метода во время обучения.

7. Оценки

- Статья — 8 (Top 50% of accepted NeurIPS papers. A very good submission; a clear accept. I vote and argue for accepting this submission)
- Уверенность — 4 (You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work)