

Дифференцируемый алгоритм поиска архитектуры модели с контролем её сложности

К. Д. Яковлев, О. С. Гребенькова, О. Ю. Бахтеев

iakovlev.kd@phystech.edu; grebenkova.os@phystech.edu; bakhteev@phystech.edu

В работе исследуется задача построения модели глубокого обучения. Предлагается метод поиска архитектуры модели, позволяющий контролировать её сложность с небольшими вычислительными затратами. Под сложностью модели понимается минимальная длина описания, минимальное количество информации, которое требуется для передачи информации о модели и о выборке. В основе метода лежит дифференцируемый алгоритм поиска архитектуры модели (DARTS). Контроль сложности параметров производится гиперсетью. Предлагается использовать гиперсеть в качестве функции релаксации. Предложенный метод позволяет контролировать сложность модели в процессе поиска архитектуры.

Ключевые слова:

1 Введение

В данной работе рассматривается оптимизация модели глубокого обучения с контролем её сложности. В качестве базового алгоритма используется DARTS. В работе [1] рассматривается проблема поиска архитектуры путем непрерывного представления архитектуры модели. Градиентные методы оптимизации позволяют использовать меньше вычислительных ресурсов. Данный алгоритм универсален для работы как со сверточными, так и с рекуррентными архитектурами нейронных сетей.

В работе [2] устойчивость алгоритма DARTS была поставлена под сомнение. Одним из источников неустойчивости является этап получения фактической дискретной архитектуры из архитектуры непрерывной смеси. На этом этапе часто наблюдается снижение качества модели. В данной работе веса модели формируются как минимизатор случайно сглаженной функции, определяемой как ожидаемая потеря в окрестности текущей архитектуры. В работе [3] предлагается использовать 0-1 функцию потерь для уменьшения расхождения между дискретной архитектурой и архитектурой непрерывной смеси.

Предлагаются альтернативные подходы к решению задачи. В работе [4] формулируется задача обучения распределению с ограничениями. Предложенный метод может быть эффективно оптимизирован и обладает теоретическими преимуществами для повышения способности к обобщению.

В работе [5] строится алгоритм поиска нейронной архитектуры с ограниченным ресурсом (RC-DARTS). К базовому алгоритму DARTS добавляются ресурсные ограничения. Для решения задачи условной оптимизации вводится алгоритм итерационной проекции.

Работе [6] исследует гиперсети. Подход заключается в использовании небольшой сети для генерации весов более крупной сети. Рассматривались два варианта использования гиперсетей: статические гиперсети для генерации весов для сверточной сети и динамические гиперсети для генерации весов рекуррентной сети.

Вычислительный эксперимент будет проводиться на выборках [7] и [8].

2 Постановка задачи

Литература

- [1] *Liu Hanxiao, Simonyan Karen, Yang Yiming*. Darts: Differentiable architecture search // CoRR, 2018. Vol. abs/1806.09055. URL: <http://arxiv.org/abs/1806.09055>.
- [2] *Chen Xiangning, Hsieh Cho-Jui*. Stabilizing differentiable architecture search via perturbation-based regularization // CoRR, 2020. Vol. abs/2002.05283.
- [3] *Chu Xiangxiang, Zhou Tianbao, 0046 Bo Zhang, Li Jixiang*. Fair darts: Eliminating unfair advantages in differentiable architecture search // CoRR, 2019. Vol. abs/1911.12126. URL: <http://arxiv.org/abs/1911.12126>.
- [4] *Chen Xiangning, Wang Ruochen, Cheng Minhao, Tang Xiaocheng, Hsieh Cho-Jui*. Drnas: Dirichlet neural architecture search // CoRR, 2020. Vol. abs/2006.10355.
- [5] *Jin Xiaojie, Wang Jiang, Slocum Joshua, 0001 Ming-Hsuan Yang, Dai Shengyang et al*. Rc-darts: Resource constrained differentiable architecture search // CoRR, 2019. Vol. abs/1912.12814. URL: <http://arxiv.org/abs/1912.12814>.
- [6] *Ha David, Dai Andrew M., Le Quoc V*. Hypernetworks // CoRR, 2016. Vol. abs/1609.09106. URL: <http://arxiv.org/abs/1609.09106>.
- [7] *LeCun Yann, Cortes Corinna*. MNIST handwritten digit database, 2010. URL: <http://yann.lecun.com/exdb/mnist/>.
- [8] *Krizhevsky Alex, Nair Vinod, Hinton Geoffrey*. Cifar-10 (canadian institute for advanced research). URL: <http://www.cs.toronto.edu/~kriz/cifar.html>.

Received February 25, 2021