

Автоматическое определение релевантности параметров нейросети

Таранов¹ С.К. Бахтеев¹ О.Ю. Стрижов^{1,2} В.В.

taranov.sk@phystech.edu; bakhteev@phystech.edu; strijov@phystech.edu

¹Московский физико-технический институт

²Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН

В данной работе исследуется выбор оптимальной структуры нейронной сети. Модели нейронных сетей зачастую содержат большое число обучаемых параметров, предполагается, что их число можно снизить с сохранением точности прогноза. Предлагается метод, корректирующий модель в процессе обучения на основе идеи представления сети в виде графа, рёбра которого являются примитивными функциями, а вершины — промежуточными представлениями выборки, полученные под действием этих функций. Для решения задачи оптимизации предлагается проводить релаксацию структуры нейросети, так чтобы соответствующая модель удовлетворяла требованиям точности для данной задачи. Также проводятся численные эксперименты на выборках данных Boston, MNIST, CIFAR-10.

Ключевые слова: *нейронные сети, оптимизация гиперпараметров, релаксация графа.*

Введение

Решается задача создания и анализа эффективного алгоритма построения нейронных сетей оптимальной сложности. Нейронные сети, очень распространённые в текущее время, обладают большой вычислительной сложностью и большим количеством обучаемых параметров. Это усложняет их обучение и использование, особенно на устройствах с ограниченными ресурсами. Поэтому для уменьшения сложности и размеров используемых моделей ведутся активные исследования в области оптимизации структур нейронных сетей.

В данной работе оптимальными будут считаться структуры, обладающие наименьшим числом гиперпараметров и обучаемых параметров, при условии сохранения достаточного уровня точности и устойчивости модели. Под гиперпараметрами будем понимать те параметры, которые описывают структуры модели, например количество нейронных слоёв, количество нейронов, содержащихся в них, а также функции активации.

Существует несколько подходов к оптимизации структуры нейронных сетей, один из них — обработка уже существующих моделей, например, прореживание уже существующих избыточных моделей с последующим дообучением [1] [2] [3]. Другой подход к проблеме оптимизации заключается в проектировании новых моделей, как результат работы алгоритмов на базе обучения с подкреплением [4], эволюционных алгоритмов [5] или как результат релаксации решения оптимизационной задачи [6]. Основываясь на [6], мы предлагаем развить описанный там алгоритм сделав его точнее благодаря использованию вариационного вывода.

Проверка алгоритма будет произведена на таких выборках как MNIST [7] и CIFAR [8], будет произведено сравнение с существующими моделями сетей, как полученными в результате других оптимизационных алгоритмов, так и построенные людьми. В качестве критериев качества будут рассматриваться в первую очередь размер модели и полное время её проектирования и обучения, а также точность.

Литература

- [1] Song Han, Jeff Pool, John Tran, and William J. Dally. Learning both weights and connections for efficient neural networks, October 30 2015. Comment: Published as a conference paper at NIPS 2015.
- [2] Hao Li, Asim Kadav, Igor Durdanovic, Hanan Samet, and Hans Peter Graf. Pruning filters for efficient convnets. *CoRR*, abs/1608.08710, 2016.
- [3] Jonathan Frankle and Michael Carbin. The lottery ticket hypothesis: Training pruned neural networks. *CoRR*, abs/1803.03635, 2018.
- [4] Barret Zoph and Quoc V. Le. Neural architecture search with reinforcement learning. *CoRR*, abs/1611.01578, 2016.
- [5] Esteban Real, Sherry Moore, Andrew Selle, Saurabh Saxena, Yutaka Leon Suematsu, Quoc V. Le, and Alex Kurakin. Large-scale evolution of image classifiers. *CoRR*, abs/1703.01041, 2017.
- [6] Hanxiao Liu, Karen Simonyan, and Yiming Yang. DARTS: differentiable architecture search. *CoRR*, abs/1806.09055, 2018.
- [7] L. Deng. The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE Signal Process. Mag*, 29(6):141–142, 2012.
- [8] Alex Krizhevsky. *Learning Multiple Layers of Features from Tiny Images*, volume 8. 2009.