

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ ГЛУБОКОГО ДОВЕРИЯ

В работе приводятся техники повышения обобщающей способности и скорости обучения нейронных сетей глубокого доверия

ВВЕДЕНИЕ

При обучении нейронных сетей глубокого доверия неизбежно возникает вопрос улучшения показателей их эффективности. К таким показателям можно отнести обобщающую способность нейронной сети и скорость выполнения обучения. Отметим, что оба показателя тесно взаимосвязаны и зачастую трудно сказать, на какую характеристику влияет тот или иной подход. Много зависит от решаемой задачи.

I. ТЕХНИКИ ПОВЫШЕНИЯ ОБОБЩАЮЩЕЙ СПОСОБНОСТИ

Повышение обобщающей способности достигается применением следующих методов: регуляризация (dropout, L1 и L2-регуляризации, искусственное расширение обучающей выборки – elastic distortions), правильный выбор целевой минимизируемой функции – mean squared error (MSE), cross-entropy (CE). Перечисленные способы актуальны как для поверхностных моделей, так и для глубоких. Для нейронных сетей глубокого доверия особое значение приобретает подход, получивший название предобучение. Предобучение – техника предварительной инициализации весов нейронной сети глубокого доверия. В настоящее время используется два подхода к предобучению: предобучение с помощью ограниченных машин Больцмана и с помощью автоассоциативных нейронных сетей. Техника предобучения с помощью ограниченных машин Больцмана была предложена Дж. Хинтоном в 2006 году [1]. Нами была предложена техника предобучения с помощью ограниченных машин Больцмана, основанная на минимизации ошибки реконструкции образа на видимом и скрытом сло-

ях [2]. Мощной альтернативой (а в некоторых случаях совмещаемым подходом) предобучению выступает использование специальных функций активации, например, ReLU (rectified linear unit), Maxout, noisy activation function и их варианты. ReLU обеспечивает устранение проблемы исчезающего градиента, тем самым обеспечивая корректную настройку весовых коэффициентов первых слоев нейронной сети.

II. ТЕХНИКИ УСКОРЕНИЯ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ

Способы ускорения обучения можно условно разделить на две категории: аппаратные и программные. К программным относятся применение моментного параметра, адаптивный шаг обучения и обучение мини-батчами. К аппаратным – использование массивно-параллельных вычислений (например, вычисления на видеокартах nVidia - технология CUDA)).

ЗАКЛЮЧЕНИЕ

Нами был выполнен обзор основных методов повышения эффективности обучения нейронных сетей глубокого доверия. Предложен эффективный подход к предобучению подобных сетей.

Список литературы

1. Hinton, G. A fast learning algorithm for deep belief nets / G. Hinton, S. Osindero, Y. Teh // Neural Computation, 18. – 2006. – pp. 1527–1554.
2. Golovko, V. A Learning Technique for Deep Belief Neural Networks / V. Golovko, A. Kroshchanka, U. Rubanau, S. Jankowski // in book Neural Networks and Artificial Intelligence. – Springer, 2014. – Vol. 440. Communication in Computer and Information Science. – P. 136–146.

Крощенко Александр Александрович, аспирант 3 года обучения факультета электронно-информационных систем Брестского государственного технического университета, kroschenko@gmail.com.

Научный руководитель: Головко Владимир Адамович, заведующий кафедрой интеллектуальных информационных технологий Брестского государственного технического университета, доктор технических наук, профессор, gva@bstu.by