

АДАПТАЦИЯ АКТИВАЦИОННЫХ ФУНКЦИЙ В ГЛУБОКИХ НЕЙРОННЫХ СЕТЯХ

Слепанская В.Д., Дейнеко А.А., Бодянский Е.В.

*Харьковский национальный университет радиоэлектроники,
г. Харьков*

В настоящее время искусственные нейронные сети (ANN) получили широкое распространение для решения широкого класса практических задач [1, 2], связанных с обработкой информации, благодаря своим универсальным аппроксимирующим возможностям и способности обучаться в процессе анализа данных. Существенно более высокого качества результатов можно добиться с помощью глубоких нейронных сетей (DNN), численная реализация которых наталкивается на вычислительные трудности, порождаемые эффектами «исчезающего» и «взрывающегося» градиента. В связи с этим, в DNN, вместо классических сигмоидальных активационных функций, обычно используются функции семейства ReLU (Rectified Linear Unit family), в которое кроме собственно ReLU входят Leaky ReLU, Parametric ReLU (PReLU), Exponential ReLU, Noisy ReLU и др. Параметры этих функций выбираются из сугубо эмпирических соображений и не изменяются в процессе обучения. DNN с этими функциями обеспечивают кусочно-линейную аппроксимацию, а сами эти функции не удовлетворяют ни одной из аппроксимационных теорем, связанных с ANN. В связи с этим для получения результата высокого качества DNN могут содержать большое количество скрытых слоев, что естественно усложняет процесс их настройки и требует больших объемов обучающих выборок.

В докладе предлагается использовать в DNN вместо фиксированных активационных функций адаптивную активационную функцию (AdPReLU), параметры которой настраиваются в процессе обучения. Таким образом, каждый нейрон сети содержит кроме набора настраиваемых синаптических весов еще два уточняемых в процессе обучения параметра кусочно-линейной активационной функции. Для настройки каждого нейрона введено модифицированное δ -правило обучения (обратное распространение для каждого нейрона) и оптимизированная процедура обратного распространения для DNN в целом.

Литература:

1. Zhernova P. Kernel fuzzy clustering of data streams based on the ensemble of neural networks / P. Zhernova, Y. Bodyanskiy // Innovative Technologies and Scientific Solutions for Industries. – 2018. – No. 4 (6). – P. 42-49. 2. Shmelov Y. Research of classification method of TV3-117 engine rating operations based on neural network technologies / Y. Shmelov, S. Vladov, O. Kryshan, S. Gvozdk, L. Chyzhova // Innovative Technologies and Scientific Solutions for Industries. – 2018. – No. 4 (6). – P. 93-102.