

1. Пролиско, Е.Е. Возможности и перспективы беспилотного городского общественного транспорта / Е.Е. Пролиско, В.Н. Шуть // Математические методы в технике и технологиях: сб. тр. межд. науч. конф. т. 9. – С-Пб : из-во Политехнического университета, 2018. – С. 16-23.

2. Сатунин, С.В. Применение комбинаторных аукционов для планирования маршрутов в моделировании задачи «транспорт по запросу» / С.В. Сатунин // Бизнес-Информатика. – 2009. – № 4 (10). – С. 3-9.

Ю.В. Савицкий

Беларусь, Брест, БрГТУ

НЕКОТОРЫЕ АСПЕКТЫ ОБУЧЕНИЯ НЕЙРОННЫХ ЭЛЕМЕНТОВ С СИГМОИДНОЙ ФУНКЦИЕЙ АКТИВАЦИИ В АРХИТЕКТУРАХ НЕЙРОННЫХ СЕТЕЙ

В настоящее время наблюдается стремительный рост исследований в направлении искусственных нейронных сетей (НС), которое является приоритетным в области искусственного интеллекта [1, 2]. При этом перспективным считается подход к предобучению (pre-training) глубоких НС не только с помощью ограниченной машины Больцмана (RBM), но и с применением нейросетевых автоэнкодеров (НА). Каждый такой НА представляет собой трехслойный персептрон архитектуры $N \rightarrow M \rightarrow N$, где параметр N соответствует количеству входов текущего предобучаемого слоя, M – количеству нейронов указанного слоя глубокой НС. Последовательное (начиная с входного слоя глубокой НС) обучение совокупности таких НА на входной обучающей выборке позволяет получить наборы весовых коэффициентов для финальной настройки синаптических связей всей сети (fine-tuning). При этом, для обучения как НА, так и глубоких НС, как правило, применяется алгоритм обратного распространения ошибки (Back Propagation Error, BPE) [3, 4], имеющий ряд принципиальных недостатков, ограничивающих его применение в практических задачах.

Определение и анализ исследуемой архитектуры НС. Объектом исследований в рамках данной работы является многослойная НС, состоящая из совокупности слоев нейроэлементов (НЭ), связанных между собой по-слоино синаптическими связями [4]. НЭ j слоя l осуществляет функцию преобразования некоторого вектора входных сигналов $Y^{[l-1]}$ НЭ предыдущего слоя в выходную активность $y_j^{[l]}$ по следующему правилу:

$$\begin{cases} S_j^{[l]} = \sum_{i=1}^{N^{[l-1]}} y_i^{[l-1]} w_{ij}^{[l]} - w_{bj}^{[l]}, \\ y_j^{[l]} = g^{[l]}(S_j^{[l]}), j = 1, \dots, N^{[l]}, \end{cases} \quad (1)$$

где $S_j^{[l]}$ - взвешенная сумма входных активностей НЭ j , находящегося в слое l ; $w_{ij}^{[l]}$ - значение синаптического веса i -го входа НЭ; $w_{bj}^{[l]}$ - значение порога активационной функции $g^{[l]}(S_j^{[l]})$ НЭ; $N^{[l-1]}, N^{[l]}$ - соответственно количество входов НЭ слоя l и количество НЭ данного слоя.

Метод точного обучения нейронных элементов с сигмоидной функцией активации. Главная идея разработанной методики точного обучения нейроэлементов с сигмоидной функцией активации сформулирована и обоснована в следующей теореме и ее строгом доказательстве (в работе не приводится).

Теорема. Правила модификации синаптических связей НЭ j , находящегося в слое L , с сигмоидной функцией активации $g^{[L]}$, минимизирующие среднеквадратичную ошибку $E_j^p(t) = 1/2(y_j^{[L],p}(t) - D_j^p)^2$ данного НЭ для эталона p на итерации обучения t , определяются следующим образом:

$$\begin{cases} w_{ij}^{[L]}(t+1) = w_{ij}^{[L]}(t) - \frac{S_j^{[L],p}(t) - \ln\left(\frac{D_j^p}{1-D_j^p}\right)}{1 + \sum_{k=1}^{N^{[L-1]}} (y_k^{[L-1],p})^2} y_i^{[L-1],p}(t), \\ w_{bj}^{[L]}(t+1) = w_{bj}^{[L]}(t) + \frac{S_j^{[L],p}(t) - \ln\left(\frac{D_j^p}{1-D_j^p}\right)}{1 + \sum_{k=1}^{N^{[L-1]}} (y_k^{[L-1],p})^2}, \end{cases} \quad (2)$$

где D_j^p — эталонное выходное значение j -го НЭ.

Таким образом, использование в алгоритме обучения правил (2) вместо правил стандартного ВРЕ обеспечивает для текущего НЭ и текущего эталона на текущей итерации нулевую ошибку обучения. Это создает предпосылки для применения данной методики в алгоритмах точного обучения НС различных архитектур, в том числе глубоких НС.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Vladimir Golovko, Yury Savitsky . New Approach of the Recurrent Neural Network Training // Proc. of the Int. Conf. on Neur. Netw. and Artif. Intell. ICNNAI'99, 12-15 October 1999. – Brest, Belarus, 1999. – P. 32-35.
2. Golovko, V. Neural Networks for Signal Processing in Measurement Analysis and Industrial Applications: the Case of Chaotic Signal Processing / V. Golovko, Y. Savitsky, N. Maniakov // chapter of NATO book «Neural networks for instrumentation, measurement and related industrial applications». – Amsterdam: IOS Press, 2003. – P. 119-143.

3. Y. Savitsky, A. Savitsky. Technique of Learning Rate Initialization for Efficient Training of MLP: Using for Computing of Lyapunov Exponents // Proc. of Int. Conf. IDT'2015, 2015 Zilina, Slovakia. – P. 287-290.

4. Савицкий Ю.В., Хвещук В.И., Савицкий А.Ю. Модифицированный алгоритм ВРЕ для адаптивного обучения сигмоидальных нейронов в архитектуре многослойной нейронной сети // Вестник БрГТУ. – Брест. – 2015. – №5(95): физика, математика, информатика. – С.38-42.