

ВЕКТОРНЫЙ АНАЛИЗ И ДЕКОМПОЗИЦИЯ – ПУТИ ПОНИМАНИЯ АЛГОРИТМОВ РАБОТЫ НЕЙРОСЕТЕЙ

Смолин В.С., Шен Х.¹

ИПМ им. М.В. Келдыша РАН, 125047, Москва, Миусская пл., д.4, +7 499 978-13-14,
smolin@keldysh.ru

¹МГУ, Факультет космических исследований, 119991, Москва, ГСП-1, 1-52, Ленинские
горы, +7 903 170-86-66, haoyuishen@yandex.ru

Успехи нейросетевых алгоритмов дают надежду, что все или почти все «интеллектуальные» задачи смогут решать машины. Хотя издаётся много учебников по глубокому обучению и генеративным нейросетям (например, [1]), практически все согласны с утверждением: «Мы не понимаем, как работают нейросети алгоритмы!».

Причины такой парадоксальной ситуации состоят в том, что при глубоком обучении сигналы и «ошибки» проходят через такое количество слоев нейросетей, что человек не может отследить, как осуществляются преобразования. Чудо состоит в том, что при обработке сложных сигналов (видеонаблюдение, преобразование речи) огромных объемов данных, используемых для обучения нейросетей теоретически недостаточно для выявления свойств сложной среды, формирующей входные сигналы. Объяснить это можно предположением, что в нейросетях формируется декомпозиция.

Исследования направлены на вектора активности элементов в пространствах латентных слоев. Матрица приращений весов входных связей ΔW^m слоя m формируется умножением вектор-столбцов $\vec{\delta}^m = \{\delta_j^m = \frac{\partial E}{\partial a_j^m}\}$ («ошибок» активности элементов слоя m) и $\vec{o}^{m-1} = \{o_i^{m-1}\}$ (выходной активности слоя $m-1$): $\Delta W^m = -\alpha \vec{\delta}^m \vec{o}^{m-1^T}$, $0 < \alpha \ll 1$. Приращения активности за счёт изменения матрицы W^m , $\Delta A^m = -\alpha \vec{\delta}^m \vec{o}^{m-1^2}$ и «ошибок» $\Delta \vec{\delta}^{m-1} = -\alpha \vec{o}^{m-1} \vec{\delta}^{m^2}$ в слое $m-1$ зависят только от векторов «ошибок» $\vec{\delta}^m$ и выходной активности \vec{o}^{m-1^2} , поскольку вторые члены в формулах произведения стоят как квадраты и не влияют на направления ΔA^m и $\Delta \vec{\delta}^{m-1}$.

Исследование векторов активности в многомерных пространствах латентных слоев позволяет по новому взглянуть на оптимизацию процессов декомпозиции сигналов и улучшение результативности нейросетевых алгоритмов [2], что открывает новые горизонты применения нейронных сетей в различных областях.

Литература

1. Фостер Д. Генеративное глубокое обучение: Как не мы рисуем картины, пишем романы и музыку / 2-е изд. – Sprint Book, 2024. – 448 с.
2. Журавлёв Д.В., Смолин В.С. Проектирование структуры нейросетей для AGI // Проектирование будущего. Проблемы цифровой реальности: труды 7-й Международной конференции (15-17 февраля 2024 г., Москва). — М.: ИПМ им. М.В.Келдыша, 2024. — С. 125- 143. <https://doi.org/10.20948/future-2024-2-3>