

Обратное распространение ошибки

Подготовили студенты гр. 5030102/20101:
Глаголев Иван, Бухаров Марк, Крупица Сергей

Что такое Backpropagation

- Backpropagation — фундаментальный метод обучения нейросетей.
- После прямого прохода вычисляется ошибка; затем ошибка «распространяется назад» и корректируются веса и смещения.
- Необходим для обучения многослойных сетей.

История возникновения

- 1960-е — ранние идеи вычисления градиентов.
- 1974 — ранний вариант алгоритма.
- 1986 — Rumelhart, Hinton, Williams: популяризация метода.
- Backprop стал основой современного глубокого обучения.

Структура простой нейронной сети

- Рассматривается сеть с входным, двумя скрытыми и выходным слоями.
- Входы — данные.
- В скрытых слоях: взвешенная сумма + функция активации.

Прямой проход (Forward pass)

1. Данные проходят через все слои.
2. На каждом нейроне: $\sum_i (w_i x_i) + b \longrightarrow f(\cdot)$
3. Получаем предсказание сети.

Вычисление ошибки

- После получения выхода — сравнение с истинным значением (target).
- Ошибка — разница между предсказанием и истинным значением.
- Используются MSE или другие функции потерь.

Обратное распространение ошибки

- Ошибка проходит назад через сеть.
- Цель — вычислить вклад каждого веса в итоговую ошибку.
- Используется правило цепочки (chain rule).

Формулы (упрощённо)

$$y = f\left(\sum_i w_i x_i + b\right)$$

$$\delta_{\text{out}} = \frac{\partial L}{\partial y} \cdot f'(z)$$

$$\delta_{\text{hidden}} = f'(z_{\text{hidden}}) \cdot \sum_j w_j \delta_{\text{next}}^{(j)}$$

$$\frac{\partial L}{\partial W} = \delta \cdot a_{\text{prev}}^{\top}$$

Ограничения и тонкости

- Функции активации должны быть дифференцируемыми.
- Проблемы: исчезающие и взрывающиеся градиенты.
- Ловушки локальных минимумов.
- Используются техники стабилизации обучения.

Заключение

- Backpropagation – базовый метод обучения нейросетей.
- Он делает возможным глубокое обучение.
- Несмотря на ограничения, остаётся одним из главных алгоритмов при обучении.

Источники

- Статья OTUS/Habr: «Знакомимся с методом обратного распространения ошибки».
- CS231n (Stanford) – Backpropagation Notes
- Michael Nielsen – Neural Networks and Deep Learning (2015)