# Исследование метода обратного распространения ошибки для обучения нейронной сети

Изображение выглядит как текст, снимок экрана, Шрифт, число

Содержимое, созданное искусственным интеллектом, может быть неверным.

Изображение выглядит как снимок экрана, текст, диаграмма, Шрифт

Содержимое, созданное искусственным интеллектом, может быть неверным.

Шаг 1. Подготовка данных

Так как мы решаем задачу линейной регрессии, и для построения прямой нужно только две точки, то для обучения можно взять только два примера — две любые строки из таблицы.

|  |  |
| --- | --- |
| x | Y |
| 0 | 32 |
| 1 | 33,8 |

Шаг 2. Обучение нейронной сети

Выполним прямой и обратный проход с одним обучающим примером (одна строчка из таблицы).

Один шаг обучения будет состоять из следующих действий:

--- Обучение на 1 обучающем примере ---

1. Прямой проход для 1 примера

2. Обратный проход для 1 примера

3. Обновление весов по результатам обратного прохода

--- Конец обучения на 1 примере –

Значение скорости обучения (learning rate, обозначается η) возьмём равным 0,01.

Начальные параметры для нейронной сети возьмём из таблицы:

|  |  |  |
| --- | --- | --- |
| Вариант | W1 | b |
| 23 | 40 | 11 |

1. **Прямой проход для 1 примера**

Используем линейную функцию активации f(S) = S, тогда на выходе выходного нейрона получаем 11.

Абсолютная ошибка сети равна

Квадратичная ошибка сети равна

1. **Обратный проход для 1 примера**

Рассчитаем производные от веса входного нейрона и от смещения выходного нейрона для того, чтобы посмотреть, как они влияют на ошибку:

1. **Обновление весов по результатам обратного прохода**

Для обновления весов воспользуемся формулой

* Поскольку результат производной равен 0, это означает, что если увеличивается на 1, то общая ошибка увеличивается на 0
* Поскольку результат производной отрицателен, это означает, что если увеличивается на 1, то общая ошибка уменьшается на 42, поэтому для уменьшения ошибки мы должны прибавлять производную.

Обновляем веса:

Заполним таблицу по заданию:

|  |  |  |
| --- | --- | --- |
| Итерация обучения | Параметры нейронной сети | |
|  |  |
| Начальное значение (до обучения) | 40 | 11 |
| После обучения на 1 примере | 40 | 11,21 |

Далее завершим первую эпоху обучения (с ещё одной строчкой таблицы) и выполним вторую (с теми же двумя строчками таблицы). Одна эпоха обучения будет состоять из следующих шагов:

--- Начало 1 эпохи ---

--- Обучение на 1 обучающем примере ---

1. Прямой проход для 1 примера

2. Обратный проход для 1 примера

3. Обновление весов по результатам обратного прохода

--- Конец обучения на 1 примере ---

--- Обучение на 2 обучающем примере ---

1. Прямой проход для 2 примера

2. Обратный проход для 2 примера

3. Обновление весов по результатам обратного прохода

--- Конец обучения на 2 примере ---

--- Конец 1 эпохи ---

Таких эпох будет две.

Завершим первую эпоху:

1. **Прямой проход для 2 примера**

Используем линейную функцию активации f(S) = S, тогда на выходе выходного нейрона получаем 51,21.

Абсолютная ошибка сети равна

Квадратичная ошибка сети равна

1. **Обратный проход для 2 примера**

Рассчитаем производные от веса входного нейрона и от смещения выходного нейрона для того, чтобы посмотреть, как они влияют на ошибку:

1. **Обновление весов по результатам обратного прохода**

Для обновления весов воспользуемся формулой

* Поскольку результат производной больше 0, это означает, что если увеличивается на 1, то общая ошибка увеличивается на 33,56
* Поскольку результат производной больше 0, это означает, что если увеличивается на 1, то общая ошибка уменьшается на 33,56, поэтому для уменьшения ошибки мы должны отнимать производную.

Обновляем веса:

Мы завершили первую эпоху и получили следующие параметры:

|  |  |
| --- | --- |
|  |  |
|  |  |

**Проделаем те же самые действия для второй эпохи**

1. **Прямой проход для 1 примера**

Используем линейную функцию активации f(S) = S, тогда на выходе выходного нейрона получаем 11,0359.

Абсолютная ошибка сети равна

Квадратичная ошибка сети равна

1. **Обратный проход для 1 примера**

Рассчитаем производные от веса входного нейрона и от смещения выходного нейрона для того, чтобы посмотреть, как они влияют на ошибку:

1. **Обновление весов по результатам обратного прохода**

Для обновления весов воспользуемся формулой

* Поскольку результат производной равен 0, это означает, что если увеличивается на 1, то общая ошибка увеличивается на 0
* Поскольку результат производной отрицателен, это означает, что если увеличивается на 1, то общая ошибка уменьшается на , поэтому для уменьшения ошибки мы должны прибавлять производную.

Обновляем веса:

Далее завершим вторую эпоху обучения (с ещё одной строчкой таблицы)

**Прямой проход для 2 примера**

Используем линейную функцию активации f(S) = S, тогда на выходе выходного нейрона получаем 51,071441.

Абсолютная ошибка сети равна

Квадратичная ошибка сети равна

1. **Обратный проход для 2 примера**

Рассчитаем производные от веса входного нейрона и от смещения выходного нейрона для того, чтобы посмотреть, как они влияют на ошибку:

1. **Обновление весов по результатам обратного прохода**

Для обновления весов воспользуемся формулой

* Поскольку результат производной больше 0, это означает, что если увеличивается на 1, то общая ошибка увеличивается на 31,347376
* Поскольку результат производной больше 0, это означает, что если увеличивается на 1, то общая ошибка уменьшается на 33,56, поэтому для уменьшения ошибки мы должны отнимать производную.

Обновляем веса:

По результатам обучения заполним таблицу по заданию

|  |  |  |  |
| --- | --- | --- | --- |
| Итерация обучения | | Параметры нейронной сети | |
|  |  |
| Начальные значения (до обучения) | | 40 | 11 |
| 1 эпоха | После обучения на 1 примере | 40 | 11,21 |
| После обучения на 2 примере | 39,8259 | 11,0359 |
| 2 эпоха | После обучения на 1 примере | 39,8259 | 11,245541 |
| После обучения на 2 примере | 39,65318559 | 11,07282659 |

Предсказания нейронной сети стали лучше после обучения, но полученная точность ещё недостаточна.