

Автор:

Дмитриев Егор Андреевич

студент

ФГАОУ ВО «Самарский национальный

исследовательский университет

им. академика С.П. Королева»

г. Самара, Самарская область

АЛГОРИТМ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ МЕТОДОМ ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

Аннотация: в данной работе рассматривается алгоритм обучения с учителем для нейронной сети – метод обратного распространения ошибки.

Ключевые слова: градиент, целевая функция, градиентный спуск.

Основные определения

Определение 1. Градиент – вектор, который указывает на направление максимального приращения функции.

Определение 2. Целевая функция – функция от нескольких переменных, значение которой нужно минимизировать или максимизировать в зависимости от рассматриваемой задачи, подбирая значения переменной.

Определение 3. Градиентный спуск – метод для нахождения оптимальных значений аргументов, которые минимизируют или максимизируют целевую функцию, путем изменения значений аргументов по направлению антиградиента или градиента.

Введение

В последнее время нейронные сети занимают лидирующие позиции в машинном обучении, решая самые разнообразные задачи. Самым сложным и крайне важным этапом является обучение нейронной сети. Метод обратного распространения ошибки является самым эффективным методом для решения задачи обучения.

Нейронные сети

Модель искусственных нейронных сетей есть математическая модель нейронной сети человека. Модель искусственной нейронной сети описывает поведение биологической, но последняя гораздо сложнее и состоит из большего количества нейронов.

Нейронные сети – это соответствие из n -мерного пространства, в пространство размерности m . На рис. 1 представлена простейшая модель нейрона – персептрон.

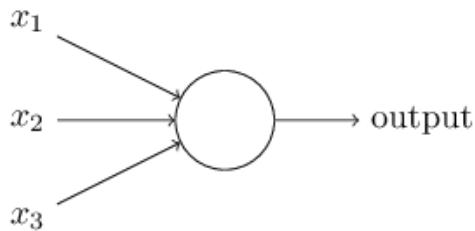


Рис. 1. Модель персептрана

$$\text{output} = \begin{cases} 0, & \text{if } w \cdot x + b \leq 0 \\ 1, & \text{if } w \cdot x + b > 0 \end{cases}$$

Дальнейшее развитие персептрана – сигмоид. Выход сигмоида – это значения функции $Q(\sum_j w_j x_j) = \frac{1}{1+\exp(-\sum_j w_j x_j)}$. Многослойная нейронная состоит из слоев сигмоидов (рис. 2).

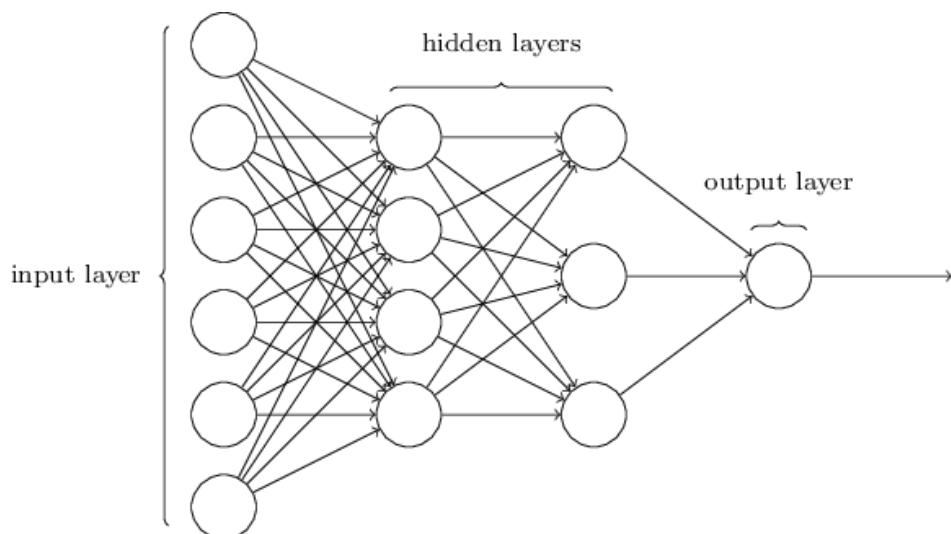


Рис. 2. Многослойная нейронная сеть

Обучение нейронной сети

Описание алгоритма обучения нейронной сети:

1. На начальном шаге происходит инициализация весов нейронной сети. На практике используется нормальное распределение с математическим ожиданием равным 0 и среднеквадратичным отклонением равным 1.

2. На входной слой подаётся вектор и определяется значение выхода нейронной сети.

3. Определяется величина ошибки для выходного слоя:

$$\delta_k = -a_k(1-a_k)(y_k - a_k).$$

4. Затем определяется ошибка для внутренних слоев

$$\delta_j = -a_j(1-a_j) \sum_k \delta_k w_{jk}.$$

5. Происходит изменения $w' = w - \eta \delta_j a_i$, $b' = b - \eta \delta_j$ где i – индекс родительского слоя, если алгоритм обучения не закончен, происходит переход на шаг 2.

6. Алгоритм завершает работу, значение ошибки перестает уменьшаться.

Так как значения ошибок нейронов предшествующих слоёв зависят от значений следующих слоев, то алгоритм называется метод обратного распространения ошибок.

Заключение

В данной работе рассмотрены модели искусственных нейронных сетей и метод обучения нейронной сети – обратное распространение ошибок.

Список литературы

1. Хайкин С. Нейронные сети. – Издательский дом Вильям, 2008.