

ТЕХНИЧЕСКИЕ НАУКИ

Татьянкин Виталий Михайлович

ст. преподаватель

ФГБОУ ВПО Югорский государственный университет

г. Ханты-Мансийск, ХМАО-Югра

МОДИФИЦИРОВАННЫЙ АЛГОРИТМ ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

Аннотация: в статье предложен модифицированный алгоритм обучения многослойных нейронных сетей. Использование предложенного алгоритма позволит уменьшить влияние субъективного мнения, что в свою очередь повысит достоверность полученных результатов.

Ключевые слова: многослойные нейронные сети, алгоритм обратного распространения ошибки.

Введение

Одним из эффективных способов обучения многослойных нейронных сетей является алгоритм обратного распространения ошибки [1,2]. Особенность рассматриваемого алгоритма заключается в том, что он функционирует, пока ошибка обучения не станет меньше желаемой. На практике выбор желаемой ошибки, очень ответственный момент и требует от пользователя высокого уровня компетенции [3]. Что в свою очередь, влияет на исход обучения, которое определяет достоверность использования многослойной нейронной сети.

Алгоритм обратного распространения ошибки.

Для обучения многослойной нейронной сети применяется алгоритм обратного распространения ошибки [1].

1. Задаётся шаг обучения $alf \in (0:1)$ и среднеквадратичная ошибка нейронной сети $Error$.

2. Случайным образом инициализируются весовые коэффициенты и пороговые значения нейронной сети.

3. Последовательно подаются вектора данных $\{x_1 \dots x_n\}^j$ на вход нейронной сети, при этом для каждого входного вектора выполняются следующие действия:

- вычисляется выходная активность всех нейронных элементов сети:

$$y_j = F(\sum_{i=1}^n w_{ij} * y_i - T_j), \quad (1)$$

где индекс j характеризует нейроны следующего слоя по отношению к слою i ;

- определяется ошибка нейронных элементов сети для выходного слоя:

$$Es_j = Y_j - t_j, \quad (2)$$

и скрытого:

$$Es_j = \sum_{i=1}^n Es_i * F'(S_i) * w_{ji}, \quad (3)$$

где индекс j характеризует нейроны следующего слоя по отношению к слою i , t_j – эталонное значение;

– изменяются весовые коэффициенты и пороги нейронных элементов по следующим правилам:

$$w_{ij}(t + 1) = w_{ij}(t) - alf * Es_j * F'(S_j) * y_i \quad (4)$$

$$T_j(t + 1) = T_j(t) + alf * Es_j * F'(S_j) \quad (5)$$

4. Вычисляется суммарная среднеквадратическая ошибка нейронной сети:

$$E = \frac{1}{2} \sum_{k=1}^L \sum_j (y_j^k - t_j^k)^2, \quad (6)$$

где L – количество обучающих векторов.

Если $E > Error$, то переходим к пункту 3. В противоположном случае алгоритм обратного распространения ошибки заканчивается.

Анализ приведённого выше алгоритма показывает, что при выборе $Error$ слишком маленькой алгоритм может не завершиться, а при выборе $Error$ слишком большой, мы можем обучить многослойную нейронную сеть, которая будет практически не применимой в силу неадекватных и недостоверных результатов.

Модифицированный алгоритм обратного распространения ошибки.

Одним из основных ресурсов, которым мы ограничены при обучении многослойной нейронной сети является время. В результате предлагается модифицированный алгоритм обратного распространения ошибки, функционирование которого определяется временем. Модифицированный алгоритм состоит из 9 шагов:

1. Задаётся шаг обучения $alf \in (0: 1)$ и время работы алгоритма t .
2. Случайным образом инициализируются весовые коэффициенты и пороговые значение нейронной сети.
3. Последовательно подаются векторы данных $\{x_1 \dots x_n\}^j$ на вход нейронной сети, при этом для каждого входного вектора выполняются следующие действия:
 - вычисляется выходная активность всех нейронных элементов сети:

$$y_j = F(\sum_{i=1}^n w_{ij} * y_i - T_j), \quad (7)$$

где индекс j характеризует нейроны следующего слоя по отношению к слою i .

- определяется ошибка нейронных элементов сети для выходного слоя:

$$Es_j = Y_j - t_j, \quad (8)$$

и скрытого:

$$Es_j = \sum_{i=1}^n Es_i * F'(S_i) * w_{ji}, \quad (9)$$

где индекс j характеризует нейроны следующего слоя по отношению к слою i , t_j – эталонное значение.

4. Вычисляется суммарная среднеквадратическая ошибка нейронной сети:

$$E = \frac{1}{2} \sum_{k=1}^L \sum_j (y_j^k - t_j^k)^2, \quad (10)$$

где L – количество обучающих векторов.

5. Создаём переменную $Error=E$.
6. Включаем счётчик времени $t1$.
7. Если $t > t1$, переходим к пункту восемь, иначе алгоритм заканчивается.
8. Весовые коэффициенты и пороги нейронных элементов изменяются по следующим правилам:

$$w_{ij}(t+1) = w_{ij}(t) - alf * Es_j * F'(S_j) * y_i \quad (11)$$

$$T_j(t+1) = T_j(t) + alf * Es_j * F'(S_j) \quad (12)$$

9. Выполняем пункты 3 и 4.

Если получившаяся ошибка меньше $Error$, то $Error=E$, запоминаем весовые коэффициенты w и порог T , переходим к пункту 7, иначе переходим в пункт 7.

Заключение

Модифицированный алгоритм обратного распространения ошибки, позволяет пользователям без опыта работы с нейронными сетями, достаточно эффективно обучить многослойную нейронную сеть. Для специалистов в области нейронных сетей, модифицированный алгоритм обратного распространение ошибки, создаёт предпосылки для разработки алгоритмов формирования оптимальной архитектуры многослойной нейронной сети.

Список литературы

1. Головко, В.А. Нейронные сети: обучение, организация и применение. Кн. 10: Учеб. пособие для вузов / Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2000.
2. Заенцев, И.В. Нейронные сети: основные модели: Издательский дом «Воронеж», 1999. – 452 с.
3. Круглов, В.В. Нечёткая логика и искусственные нейронные сети: ФИЗМАЛИТ, 2001. – 221 с.