

ОБ ОБУЧЕНИИ НЕЙРОННЫХ СЕТЕЙ С ПОМОЩЬЮ ИЗБИРАТЕЛЬНОГО МЕТОДА МОНТЕ-КАРЛО

Мазуров М. Е.

Российский Экономический Университет им. Г.В. Плеханова

Россия, 117997, г.Москва, Стремянный пер., 36

Тел.: 8 916 190 25 15

E-mail: mazurov37@mail.ru

Обучение многослойных нейронных сетей, использующих нейроны МакКаллока - Питтса, сводится к расчету весовых коэффициентов с помощью рекуррентного преобразования $y_{i+1}^t = F^t(W_i^t y_i^t + U_i^t)$, где $y_i = (y_{i1}, \dots, y_{in})$. Матрицы W_i^t вычисляются для каждого i -го слоя нейронной сети с помощью итерационных процедур:

$$W_i^{t+1} = W_i^t + \Delta W_i^t, \quad d_m^t = \sum_{i=1}^m (y_m^t - y_0^t)^2, \quad \text{где } \Delta W_i^t \text{ - приращения весовых коэффициентов}$$

для итерации с номером t ; d_m^t - ошибка вычислений. Итерационный процесс имеет ряд не устранимых недостатков. Перечислим некоторые из них: 1. Коэффициенты весовой матрицы W могут быть либо слишком большими, либо слишком малыми, соизмеримыми с шумами, или другими не подходящими, поскольку задача определения весовых коэффициентов бесконечно значима; 2. Итерационная процедура может быть значительной по времени; 3. Нет четкой процедуры изменения огромного количества весовых коэффициентов; 4. Отсутствует процедура обоснованного выбора начальных условий и так далее.

Вместо итерационного способа определения весовых коэффициентов в данной работе предложено их нахождение с помощью метода Монте-Карло. Производится избирательное нахождение случайных чисел, учитывающее требуемый диапазон нахождения весовых коэффициентов, требуемые пороговые условия, требуемые начальные условия.

Для генерации случайных чисел было использовано программное обеспечение Матлаб 7. Для получения случайных чисел из произвольного диапазона $[a, b]$ использовался оператор $a+(b-a)*\text{rand}$. Было рассмотрено распознавание однослойной нейронной сетью Розенблюта 10 цифр 0, 1, ..., 9 на экране монитора 4x6. Общее число весовых коэффициентов равно $24 \times 10 = 240$, то есть матрица W имеет размерность 24×10 . Избирательные свойства матрицы весовых коэффициентов характеризуются матрицей $B = W \times W^T$ и иллюстрируются справа.

Предлагаемый метод позволяет избавиться от недостатков итерационного метода, изложенных в п.п. 1-4, получить несколько наборов весовых коэффициентов для выбора наиболее подходящих.

