

## Визуализация процессов обучения нейронных сетей

*Емельянова Ю. Г., Мальшевский А. А., Хачумов В. М.*

vmh@vmh.botik.ru, Scorpio@Zhukovsky.net, tajra@mail.ru

Переславль-Залесский, Институт программных систем РАН

Рассматривается задача когнитивной графической поддержки процессов настройки искусственных нейронных сетей (ИНС). Визуализация позволяет непосредственно наблюдать за ходом обучения ИНС, предоставляя в наглядном виде информацию о степени готовности и возникающих аномалиях средствами машинной графики, что позволяет пользователю оперативно принимать решение [1, 2].

### Постановка задачи

Рассмотрим постановку задачи для двух принципиально различных по структуре ИНС. Данному условию соответствуют, например, сеть с единственным выходом, реализующая логическую функцию XOR [2], и однослойный персептрон с десятью выходами, предназначенный для распознавания цифр. В первом случае используется генетический алгоритм настройки и активационная функция вида «сигмоид», а во втором — алгоритм Видроу–Хоффа и функция активации типа «линейный скачок». Требуется отобразить графически динамику процессов обучения нейронной сети. Для визуализации процессов обучения нейронной сети, прежде всего, необходима информация о весах синапсов — односторонних входных связей отдельных нейронов, которые подстраиваются в процессе обучения.

### Визуализация настройки сети XOR

Поставим в соответствие сети (Рис. 1) вектор из семи прямоугольных полутоновых областей. Примем, что белый цвет обозначает ноль, оттенки синего — отрицательные значения, оттенки красного — положительные значения коэффициента. Чем больше по модулю значение весового коэффициента, тем ярче цвет соответствующего прямоугольника. Под областью вектора расположен зеленый квадрат, который «мигает» при каждой четной итерации. Это представление позволяет сопоставить скорость роста фитнес-функции со скоростью совершения итераций; помогает заметить ситуацию, при которой веса синапсов не меняются или меняются незначительно от итерации к итерации. Значение фитнес-функции отображается в виде оранжевой полосы. Чем ярче цвет и чем больше ее длина, тем ближе значения полученного и требуемого выходов. При удачном завершении настройки на экране появляется информация о количестве совершенных итераций, таблица входов и соответствующих им выходов (Рис. 2). Интерпретация получаемых на экране цветовых

картин дана в работе [1] и в настоящей статье модифицируется следующим образом.

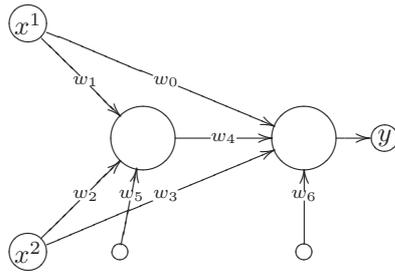


Рис. 1. ИНС XOR.

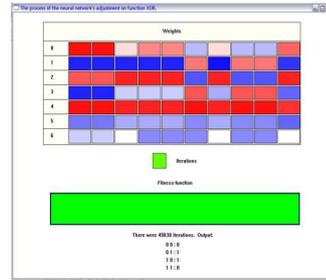


Рис. 2. Вид окна при успешной настройке сети.

1. Если значение фитнес-функции не изменяется от итерации к итерации и большинство ячеек ИНС окрашено ярким цветом, то сеть впала в «паралич».
2. Если значение фитнес-функции не меняется от итерации к итерации и большинство полос имеет некоторую промежуточную окраску, также не изменяющуюся во времени, то сеть вошла в локальный минимум.
3. Если когнитивный образ меняется от итерации к итерации, а фитнес-функция при этом практически не растет, то сеть не настраивается.

### Визуализация настройки персептрона

Рассмотрим визуализацию настройки персептрона (Рис. 3, 4), предназначенного для распознавания цифр. Графический динамический образ входа нейрона строится в виде цветного прямоугольника (полосы) с изменяющейся толщиной. Для представления знака весового коэффициента применяется цвет: красный — для положительного, синий — для отрицательного. Толщина прямоугольника служит для представления динамики абсолютных значений весов: чем толще полоса, тем быстрее изменяются значения весов. С уменьшением толщины процесс стабилизируется, когда же полосы исчезают и более не появляются — веса нейрона практически настроены. Для представления выходов нейронов выбран подход с использованием полос синего и красного цветов, а также их оттенков. Оттенки прямоугольников меняются, отображая значения выходов. Каждый нейрон сети настраивается на эталонное изображение отдельно, поэтому, наблюдая настройку сети, мы видим сначала меняющееся изображение полосы первого нейрона, затем второго при статичном первом и т.д., как это показано на Рис. 4. При этом прямоугольник, отвечаю-

щий за первый нейрон, т.е. за тот, который уже настроен, приобретает зеленый цвет. Визуализация настройки персептрона, позволяет, как и в случае ИНС XOR, определять такие явления как паралич или локальный минимум. Ситуация, когда прямоугольник, отвечающий за выход с нейрона, окрашен в яркий цвет и не меняет свой оттенок длительное время, означает, что нейрон впал в паралич. Если оба прямоугольника, относящихся к одному нейрону, во время его настройки приняли статичное промежуточное положение, а нейрон все ещё является ненастроенным, то имеем локальный минимум.

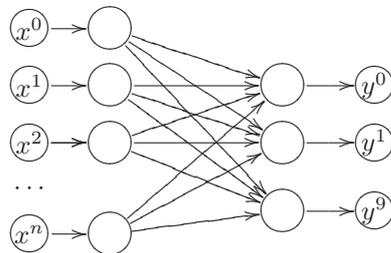


Рис. 3. Персептрон.

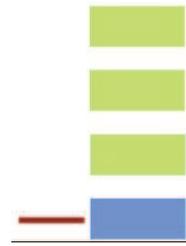


Рис. 4. Пример визуализации.

### Заключение

В настоящей работе предложены подходы к визуализации процессов обучения некоторых нейронных сетей, имеющих различное назначение и архитектуру. Разработанные графические динамические представления весов нейронов ИНС и их выходов позволяют контролировать текущее состояние сети, оценивать качество и скорость обучения. Они дают общую картину внутреннего состояния ИНС не только в процессе обучения сети, но и в процессе ее непосредственного использования.

Работа выполнена при поддержке РФФИ, проект № 06-07-89083.

### Литература

- [1] Таран А., Москаленко Ю. Визуализация процесса обучения нейронной сети. — [http://old.festu.ru/ru/structure/library/Library/science/s127/article\\_9.htm](http://old.festu.ru/ru/structure/library/Library/science/s127/article_9.htm).
- [2] Хачумов В. М. Логические элементы на нейронах // Интеллектуальные системы и компьютерные науки, Москва: Издательство механико-математического факультета МГУ, 2006. — С. 297–300.