

СИСТЕМА РАСПОЗНАВАНИЯ СИМВОЛОВ С ПОМОЩЬЮ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

Рог В. И.

Научный руководитель: Друки А. А.
Томский политехнический университет
E-mail: vir7@tpu.ru

Введение

В последнее время в мире активно развивается новое направление исследований, которое направлено на проектирование искусственных нейронных сетей. Актуальность нейронных сетей подтверждается широким спектром областей их применения. С помощью нейронных сетей можно, к примеру, выполнять распознавать различные звуковые или оптические сигналы, можно создать системы, способные управлять машиной при парковке и т. д.

Описание ИНС

Искусственная нейронная сеть (ИНС) — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. Первой такой попыткой были нейронные сети Маккалока и Питтса. После разработки алгоритмов обучения получаемые модели стали использовать в практических целях: в задачах прогнозирования, для распознавания образов, в задачах управления и др. Схема простой нейронной сети представлена на рисунке 1.

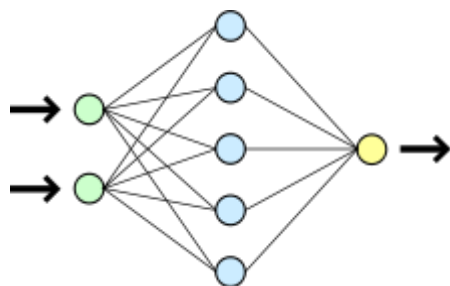


Рис. 1. Схема простой нейросети. Зелёным цветом обозначены входные нейроны, голубым — скрытые нейроны, жёлтым — выходной нейрон

ИНС представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Такие процессоры обычно довольно просты (особенно в сравнении с процессорами, используемыми в персональных компьютерах). Каждый процессор подобной сети имеет дело только с сигналами, которые он периодически получает и сигналами, которые он периодически отправляет другим процессорам. Соединённые в одну достаточно

большую сеть такие простые процессоры могут выполнять сложные задачи.

Нейронные сети не программируются, они обучаются. Это одно из главных преимуществ нейронных сетей перед алгоритмами. Технически обучение заключается в нахождении коэффициентов связи между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что в случае успешного обучения нейронная сеть сможет вернуть верный результат на основании данных, которые отсутствуют в обучающей выборке, а также неполных и/или «зашумленных» данных.

Постановка задачи

Реализовать нейронную сеть и обучить ее распознавать цифры от 0 до 9, заданные в матричном виде 5x7. Цифры представлены на рисунке 2.

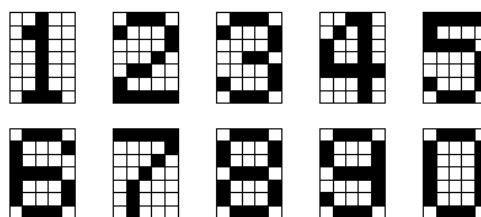


Рис. 2. Цифры в матричном виде

Разработка ИНС

Для начала определим входные и выходные данные нейронной сети. Каждая цифра записана в ячейке 5x7, значит входные данные будут состоять из 35 элементов. На выходе мы получаем набор из 10 элементов, по количеству цифр. В качестве активации будем использовать лог-сигмоидную функцию, представленную на рисунке 3.

$$f(S) = \frac{1}{1 + e^{-aS}}$$

Рис. 3. Лог-сигмоидная функция активации, где a — константа, S — взвешенная сумма входных сигналов нейрона.

Функция активации (активационная функция, функция возбуждения) — функция, вычисляющая выходной сигнал искусственного нейрона.

Сеть будет состоять из трех слоев (входной, скрытый, выходной).

Скрытые слои представляют собой нейроны, которые не имеют непосредственных входов исходных данных, а связаны только с выходами входного слоя и с входом выходного слоя. Таким образом, скрытые слои дополнительно преобразуют информацию и добавляют нелинейности в модели

Для обучение нейронной сети возьмем алгоритм обратного распространения ошибки. При использовании лог-сигмоидной функции значение ошибки для нейрона выходного слоя и нейрона, одному из скрытых слоев примут вид:

$$\delta_i = af_{i,k}(1 - f_{i,k})(f_{i,k} - y_{i,k})$$

Рис.4. Ошибка нейрона для выходного слоя

Где $y_{i,k}$ – требуемое, а $f_{i,k}$ – фактическое значение выходного сигнала k -го нейрона для i -го набора данных из обучающей выборки, $(f_{i,k}(S))'$ – значение производной активационной функции k -го нейрона для i -го набора обучающих данных.

$$\delta_i^{(q)} = af_i(1 - f_i) \sum_j w_{i,j} \delta_j^{(q+1)}$$

Рис.5. Ошибка нейрона для скрытого слоя

Где $\delta_i^{(q)}$ – ошибка i -го нейрона в слое q , $\delta_j^{(q+1)}$ – ошибка j -го нейрона в $(q+1)$ слое, w_{ij} – вес связи, соединяющей эти нейроны, $(f_{i,k}(S))'$ – значение производной активационной функции i -го нейрона слоя q .

В качестве функции ошибки ИНС будем рассматривать следующую величину:

$$E(w) = \frac{1}{2} \sum_i E_i = \frac{1}{2} \sum_{i,k} (f_{i,k} - y_{i,k}^{(T)})^2.$$

Рис.6. Функция ошибки ИНС

Для реализации алгоритма обратного распространения ошибки может быть использована следующая последовательность действий:

1. Предъявление очередного набора из обучающей выборки на вход нейронной сети.
2. Вычисление выходного сигнала сети.
3. Определение величин ошибок нейронов выходного слоя по формуле (Рис.4.).
4. Определение величин ошибок нейронов скрытых слоев по формуле (Рис.5.).
5. Однократная коррекция весов связей.
6. Если в обучающей выборке есть неиспользованные в данной эпохе наборы данных, то переход на шаг 1.
7. Подсчет ошибки сети по формуле (Рис.6.). Если ошибка меньше заданной, то конец обучения, иначе, начало новой эпохи обучения и переход на шаг 1.

Отметим, что алгоритм обратного распространения ошибки применим только для нейронных сетей, содержащих нейроны с

дифференцируемой функцией активации. Т.е. рассмотренный алгоритм не подходит для настройки сетей, построенных на нейронах с пороговыми функциями активации. Для таких сетей применяются другие алгоритмы обучения, например, дельта-правило Уидроу-Хоффа.

В настоящее время создано множество модификаций алгоритма обратного распространения ошибки. Одним из них является алгоритм, использующий инерционность обучения. В нем вес связи изменяется в соответствии с формулой на рисунке 7.

$$w_{ij}(t+1) = w_{ij}(t) - \eta \delta_j x_i + \alpha \Delta w_{ij}(t-1),$$

Рис.7. Вес связи в алгоритме, использующем инерционность обучения.

где α – коэффициент инерционности от 0 до 1. Данный алгоритм сходится в среднем в 4–5 раз быстрее стандартного, но использует больший объем памяти.

Заключение

На данном этапе разработки была поставлена задача по реализации ИНС и построен алгоритм ее работы. В рамках решаемой задачи был изучен алгоритм обратного распространения ошибки.

В результате был завершен этап проектирования ИНС, проведена подготовка и начата его реализации.

Список использованных источников

1. Применение нейросетей в распознавании изображений// Habrhabr // URL: <https://geektimes.ru/post/74326/> (дата обращения: 03.06.2016).
2. Artificial Neural Networks/Activation Functions// Wikibooks // URL: https://en.wikibooks.org/wiki/Artificial_Neural_Networks/Activation_Functions (дата обращения: 03.06.2016).
3. Хайкин С. Нейронные сети: полный курс, 2-е издание. М.: Издательский дом «Вильямс», 2006.
4. Каллан Р. Основные концепции нейронных сетей: пер. с англ. – М.: Издательский дом «Вильямс», 2001. – 288 с.
5. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002. – 344 с.
6. Распознавание образа нейросетью// Портал Ampersant // URL: <http://www.ampersant.ru/glaz/> (дата обращения: 04.06.2016).