

Курс лекций «Введение в ИИ.». Часть III.

Нейрокомпьютеринг.

Лекция 15. Персептроны

О.Г. Чанышев

Содержание

1 Персептроны и зарождение искусственных нейронных сетей	1
2 ОБУЧЕНИЕ ПЕРСЕПТРОНА	2
2.1 Алгоритм обучения персептрона	2
2.2 Дельта-правило	3

1 Персептроны и зарождение искусственных нейронных сетей

Персептроны иллюстрируют важные принципы. В силу этих причин они являются логической исходной точкой для изучения искусственных нейронных сетей.

В качестве научного предмета искусственные нейронные сети впервые заявили о себе в 40-е годы.

Первое систематическое изучение искусственных нейронных сетей было предпринято Маккаллоком и Питтсом в 1943 г. [1]. Позднее в работе [3] они исследовали сетевые парадигмы для распознавания изображений, подвергаемых сдвигам и поворотам. Простая нейронная модель, показанная на 1, использовалась в большей части их работы. Элемент Σ умножает каждый вход x на вес w и суммирует взвешенные входы. Если эта сумма больше заданного порогового значения, выход равен единице, в противном случае – нулю. Эти системы (и множество им подобных) получили название персептронов. Они состоят из одного слоя искусственных нейронов, соединенных с помощью весовых коэффициентов с множеством входов (см. 2), хотя в принципе описываются и более сложные системы.

Открытие методов обучения многослойных сетей в большей степени, чем какой-либо иной фактор, повлияло на возрождение интереса и исследовательских усилий.

Несмотря на свои ограничения персептроны широко изучались (хотя не слишком широко использовались).

Эффективность запоминания Серьезные вопросы имеются относительно эффективности запоминания информации в персептроне (или любых других нейронных сетях) по сравнению

Рис. 1: Персептронный нейрон

Рис. 2: Персептрон со многими выходами

Рис. 3: Персептронная система распознавания изображений

с обычной компьютерной памятью и методами поиска информации в ней. Например, в компьютерной памяти можно хранить все входные образы вместе с классифицирующими битами. Компьютер должен найти требуемый образ и дать его классификацию. Различные хорошо известные методы могли бы быть использованы для ускорения поиска. Если точное соответствие не найдено, то для ответа может быть использовано правило ближайшего соседа.

Число битов, необходимое для хранения этой же информации в весах персептрона, может быть значительно меньшим по сравнению с методом обычной компьютерной памяти, если образы допускают экономичную запись. Однако Минский [2] построил патологические примеры, в которых число битов, требуемых для представления весов, растет с размерностью задачи быстрее, чем экспоненциально. В этих случаях требования к памяти с ростом размерности задачи быстро становятся невыполнимыми. Если, как он предположил, эта ситуация не является исключением, то персептроны часто могут быть ограничены только малыми задачами. Насколько общими являются такие неподатливые множества образов? Это остается открытым вопросом, относящимся ко всем нейронным сетям. Поиски ответа чрезвычайно важны для исследований по нейронным сетям.

2 ОБУЧЕНИЕ ПЕРСЕПТРОНА

Алгоритм обучения персептрона может быть реализован на цифровом компьютере или другом электронном устройстве, и сеть становится в определенном смысле самоподстраивающейся. По этой причине процедуру подстройки весов обычно называют «обучением» и говорят, что сеть «обучается».

Обучение может быть с учителем или без него. Для обучения с учителем нужен «внешний» учитель, который оценивал бы поведение системы и управлял ее последующими модификациями. При обучении без учителя сеть путем самоорганизации делает требуемые изменения. Обучение персептрона является обучением с учителем.

2.1 Алгоритм обучения персептрона

Персептрон обучают, подавая множество образов по одному на его вход и подстраивая веса до тех пор, пока для всех образов не будет достигнут требуемый выход. Допустим, что входные образы нанесены на демонстрационные карты. Каждая карта разбита на квадраты и от каждого квадрата на персептрон подается вход. Если в квадрате имеется линия, то от него подается единица, в противном случае – ноль. Множество квадратов на карте задает, таким образом, множество нулей и единиц, которое и подается на входы персептрона. Цель состоит в том, чтобы научить персептрон включать индикатор при подаче на него множества входов, задающих нечетное число, и не включать в случае четного.

На 3 показана такая персептронная конфигурация. Допустим, что вектор \vec{X} является образом распознаваемой демонстрационной карты. Каждая компонента $X = (x_1, x_2, \dots, x_n)$ – умножается на соответствующую компоненту вектора весов $W = (w_1, w_2, \dots, w_n)$. Эти произведения суммируются ($NET = \sum_i x_i w_i = \vec{X}\vec{W}$). Если сумма превышает порог Θ , то выход нейрона Y

равен единице (индикатор зажигается), в противном случае он – ноль, а после нее следует пороговая операция.

Для обучения сети образ \vec{X} подается на вход и вычисляется выход Y . Если Y правилен, то ничего не меняется. Однако если выход неправилен, то веса, присоединенные к входам, усиливающим ошибочный результат, модифицируются, чтобы уменьшить ошибку.

Чтобы увидеть, как это осуществляется, допустим, что демонстрационная карта с цифрой 3 подана на вход и выход Y равен 1 (показывая нечетность). Так как это правильный ответ, то веса не изменяются. Если, однако, на вход подается карта с номером 4 и выход Y равен единице (нечетный), то веса, присоединенные к единичным входам, должны быть уменьшены, так как они стремятся дать неверный результат. Аналогично, если карта с номером 3 дает нулевой выход, то веса, присоединенные к единичным входам, должны быть увеличены, чтобы скорректировать ошибку. Этот метод обучения может быть подытожен следующим образом:

- 1. Подать входной образ и вычислить Y .
- 2а. Если выход правильный, то перейти на шаг 1;
- 2б. Если выход неправильный и равен нулю, то добавить все входы к соответствующим им весам; или
- 2в. Если выход неправильный и равен единице, то вычесть каждый вход из соответствующего ему веса.
- 3. Перейти на шаг 1.

За конечное число шагов сеть научится разделять карты на четные и нечетные. Это значит, что для всех нечетных карт выход будет больше порога, а для всех четных – меньше. Отметим, что это обучение глобально, т. е. сеть обучается на всем множестве карт. Возникает вопрос о том, как это множество должно предъявляться, чтобы минимизировать время обучения. Должны ли элементы множества предъявляться последовательно друг за другом или карты следует выбирать случайно? Несложная теория служит здесь путеводителем.

2.2 Дельта-правило

Важное обобщение алгоритма обучения персептрона, называемое дельта-правилом, переносит этот метод на непрерывные входы и выходы. Чтобы понять, как оно было получено, шаг 2 алгоритма обучения персептрона может быть сформулирован в обобщенной форме с помощью введения величины δ , которая равна разности между требуемым или целевым выходом T и реальным выходом Y

$$\delta = (T - Y). \quad (1)$$

Случай, когда $\delta = 0$, соответствует шагу 2а, когда выход правилен и в сети ничего не изменяется. Шаг 2б соответствует случаю $\delta > 0$, а шаг 2в случаю $\delta < 0$. В любом из этих случаев персептронный алгоритм обучения сохраняется, если δ умножается на величину каждого входа x_i и это произведение добавляется к соответствующему весу. С целью обобщения вводится коэффициент «скорости обучения» η , который умножается на δ_i , что позволяет управлять средней величиной изменения весов.

В алгебраической форме записи

$$\Delta_i = \eta \delta x_i, \quad (2)$$

$$w_i^{(n+1)} = w_i^{(n)} + \Delta_i, \quad (3)$$

где Δ_i – коррекция, связанная с i -м входом x_i ; $w_i^{(n+1)}$ – значение i -го после коррекции; $w_i^{(n)}$ – значение веса i -го до коррекции.

Дельта-правило модифицирует веса в соответствии с требуемым и действительным значениями выхода каждой полярности как для непрерывных, так и для бинарных входов и выходов. Эти свойства открыли множество новых приложений.

Список литературы

- [1] McCulloch W. W., Pitts W. 1943. A logical calculus of the ideas imminent in nervous activiti. Bulletin of Mathematical Biophysics 5:115-33. (Русский перевод: Маккаллоу У. С., Питтс У. Логическое исчисление идей, относящихся к нервной деятельности. Автоматы. – М.: ИЛ. – 1956.
- [2] Minsky M. L, Papert S. 1969. Perceptrons. Cambridge, MA: MIT Press. (Русский перевод: Минский М. Л., Пейперт С. Перцептроны. – М: Мир. – 1971.)
- [3] Pitts W. Mcculloch W. W. 1947. How we know universals. Bulletin of Mathematical Biophysics 9:127-47.
- [4] Rosenblatt F. 1962. Principles of Neurodynamics. New York: Spartan Books. (Русский перевод: Розенблатт Ф. Принципы нейродинамики. – М: Мир. – 1965.)
- [5] Widrow B. 1961. The speed of adaptation in adaptive control system, paper *1933- 61. American Rocket Society Guidance Control and Navigation Conference.
- [6] Widrow B. 1963. A statistical theory of adaptation. Adaptive control systems. New York: Pergamon Press.
- [7] Widrow B., Angell J. B. 1962. Reliable, trainable networks for computing and control. Aerospace Engineering 21:78-123.
- [8] Widrow B., Hoff M. E. 1960. Adaptive switching circuits. 1960 IRE WESCON Convention Record, part 4, pp. 96-104. New York: Institute of Radio Engineers.

Следующая лекция

[Лекция 16. Процедура обратного распространения](#)