

Искусственная нейронная сеть

Материал из Википедии — свободной энциклопедии

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 У. Маккалока и У. Питтса^[1]. После разработки алгоритмов обучения получаемые модели стали использовать в практических целях: в задачах прогнозирования, для распознавания образов, в задачах управления и др.

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 посылает другим процессорам. И, тем не менее, будучи соединёнными в достаточно большую сеть с управляемым взаимодействием, такие по отдельности простые процессоры вместе способны выполнять довольно сложные задачи.

- С точки зрения машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т. п.
- С математической точки зрения, обучение нейронных сетей — это многопараметрическая задача нелинейной оптимизации.
- С точки зрения кибернетики, нейронная сеть используется в задачах адаптивного управления и как алгоритмы для робототехники.
- С точки зрения развития вычислительной техники и программирования, нейронная сеть — способ решения проблемы эффективного параллелизма^[2].
- А с точки зрения искусственного интеллекта, ИНС является основой философского течения коннективизма и основным направлением в структурном подходе по изучению возможности построения (моделирования) естественного интеллекта с помощью компьютерных алгоритмов.

Нейронные сети не программируются в привычном смысле этого слова, они **обучаются**. Возможность обучения — одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей 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 сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искажённых данных.

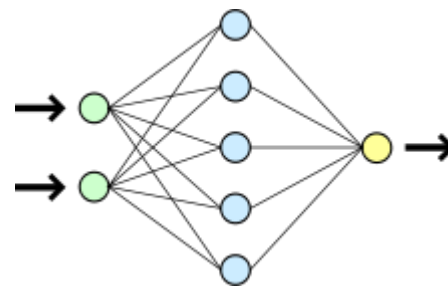


Схема простой нейросети. Зелёным цветом обозначены входные нейроны, голубым — скрытые нейроны, жёлтым — выходной нейрон

Содержание

- 1 Хронология
- 2 Известные применения
 - 2.1 Распознавание образов и классификация
 - 2.1.1 Используемые архитектуры нейросетей
 - 2.2 Принятие решений и управление

- 2.2.1 Используемые архитектуры нейросетей
 - 2.3 Кластеризация
 - 2.3.1 Используемые архитектуры нейросетей
 - 2.4 Прогнозирование
 - 2.4.1 Используемые архитектуры нейросетей
 - 2.5 Аппроксимация
 - 2.5.1 Используемые архитектуры нейросетей
 - 2.6 Сжатие данных и Ассоциативная память
 - 2.6.1 Используемые архитектуры нейросетей
 - 2.7 Анализ данных
 - 2.7.1 Используемые архитектуры нейросетей
 - 2.8 Оптимизация
 - 2.8.1 Используемые архитектуры нейросетей
- 3 Этапы решения задач
 - 3.1 Сбор данных для обучения
 - 3.2 Выбор топологии сети
 - 3.3 Экспериментальный подбор характеристик сети
 - 3.4 Экспериментальный подбор параметров обучения
 - 3.5 Обучение сети
 - 3.6 Проверка адекватности обучения
- 4 Классификация по типу входной информации
- 5 Классификация по характеру обучения
- 6 Классификация по характеру настройки синапсов
- 7 Классификация по времени передачи сигнала
- 8 Классификация по характеру связей
 - 8.1 Сети прямого распространения (Feedforward)
 - 8.2 Рекуррентные нейронные сети
 - 8.3 Радиально-базисные функции
 - 8.4 Самоорганизующиеся карты
- 9 Известные типы сетей
- 10 Отличия от машин с архитектурой фон Неймана
- 11 Примеры приложений
 - 11.1 Предсказание финансовых временных рядов
 - 11.2 Психодиагностика
 - 11.3 Хемоинформатика
 - 11.4 Нейроуправление
 - 11.5 Экономика
- 12 См. также
- 13 Примечания
- 14 Литература
- 15 Ссылки

Хронология

- 1943 — У. Маккалок и У. Питтс формализуют понятие нейронной сети в фундаментальной статье о логическом исчислении идей и нервной активности^[1]. В начале своего сотрудничества с Питтсом Н. Винер предлагает ему вакуумные лампы в качестве идеального на тот момент средства для реализации эквивалентов нейронных сетей^[3].
- 1948 — Н. Винер вместе с соратниками публикует работу о кибернетике. Основной идеей является представление сложных биологических процессов математическими моделями.
- 1949 — Д. Хебб предлагает первый алгоритм обучения.
- В 1958 Ф. Розенблат изобретает однослойный перцептрон и демонстрирует его способность решать задачи классификации^[4]. Перцептрон обрёл популярность — его используют для распознавания образов, прогнозирования погоды и т. д.; в то время казалось, что уже не за горами создание полноценного искусственного интеллекта. К моменту изобретения перцептрона

завершилось расхождение теоретических работ Маккалока с т. н. «кибернетикой» Винера; Маккалок и его последователи вышли из состава «Кибернетического клуба».

- В 1960 году Уидроу совместно со своим студентом Хоффом на основе дельта-правила (формулы Уидроу) разработали Адалин, который сразу начал использоваться для задач предсказания и адаптивного управления. Адалин был построен на базе созданных ими же (Уидроу — Хоффом) принципиально новых элементах — мемисторах^[5]. Сейчас Адалин (адаптивный сумматор) является стандартным элементом многих систем обработки сигналов^[6].
- В 1963 году в Институте проблем передачи информации АН СССР А. П. Петровым проводится подробное исследование задач «трудных» для перцептрона^[7]. Эта пионерская работа в области моделирования ИНС в СССР послужила отправной точкой для комплекса идей М. М. Бонгарда — как «сравнительно небольшой переделкой алгоритма (перцептрона) исправить его недостатки»^[8]. Работы А. П. Петрова и М. М. Бонгарда весьма способствовали тому, что в СССР первая волна эйфории по поводу ИНС была сглажена.
- В 1969 году М. Минский публикует формальное доказательство ограниченности перцептрона и показывает, что он не способен решать некоторые задачи (проблема «чётности» и «один в блоке»), связанные с инвариантностью представлений. Интерес к нейронным сетям резко падает.
- В 1972 году Т. Кохонен и Дж. Андерсон независимо предлагают новый тип нейронных сетей, способных функционировать в качестве памяти^[9].
- В 1973 году Б. В. Хакимов предлагает нелинейную модель с синапсами на основе сплайнов и внедряет её для решения задач в медицине, геологии, экологии^[10].
- 1974 — Пол Дж. Вербос^[11] и А. И. Галушкин^[12] одновременно изобретают алгоритм обратного распространения ошибки для обучения многослойных перцептронов^[уточнить]. Изобретение не привлекло особого внимания.
- 1975 — Фукусима представляет когнитрон — самоорганизующуюся сеть, предназначенную для инвариантного распознавания образов, но это достигается только при помощи запоминания практически всех состояний образа.
- 1982 — после периода забвения, интерес к нейросетям вновь возрастает. Дж. Хопфилд показал, что нейронная сеть с обратными связями может представлять собой систему, минимизирующую энергию (так называемая сеть Хопфилда). Кохоненом представлены модели сети, обучающейся без учителя (нейронная сеть Кохонена), решающей задачи кластеризации, визуализации данных (самоорганизующаяся карта Кохонена) и другие задачи предварительного анализа данных.
- 1986 — Дэвидом И. Румельхартом, Дж. Е. Хинтоном и Рональдом Дж. Вильямсом^[13] и независимо и одновременно С. И. Барцевым и В. А. Охониным (Красноярская группа)^[14] переоткрыт и существенно развит метод обратного распространения ошибки. Начался взрыв интереса к обучаемым нейронным сетям.
- 2007 Джеффри Хинтоном в университете Торонто созданы алгоритмы глубокого обучения многослойных нейронных сетей. Успех обусловлен тем, что Хинтон при обучении нижних слоев сети использовал ограниченную машину Больцмана (RBM — Restricted Boltzmann Machine). Глубокое обучение по Хинтону — это очень медленный процесс. Необходимо использовать много примеров распознаваемых образов (например, множество лиц людей на разных фонах). После обучения получается готовое быстро работающее приложение, способное решать конкретную задачу (например, осуществлять поиск лиц на изображении). Функция поиска лиц людей на сегодняшний день стала стандартной и встроена во все современные цифровые фотоаппараты. Технология глубокого обучения активно используется интернет-поисковиками при классификации картинок по содержащимся в них образам. Применяемые при распознавании искусственные нейронные сети могут иметь до 9 слоев нейронов, их обучение ведётся на миллионах изображений с отыскиваемым образом.

Известные применения

Распознавание образов и классификация

В качестве образов могут выступать различные по своей природе объекты: символы текста, изображения, образцы звуков и т. д. При обучении сети предлагаются различные образцы образов с указанием того, к какому классу они относятся. Образец, как правило, представляется как вектор

значений признаков. При этом совокупность всех признаков должна *однозначно определять класс*, к которому относится образец. В случае, если признаков недостаточно, сеть может соотнести один и тот же образец с несколькими классами, что неверно. По окончании обучения сети ей можно предъявлять неизвестные ранее образы и получать ответ о принадлежности к определённому классу.

Топология такой сети характеризуется тем, что количество нейронов в выходном слое, как правило, равно количеству определяемых классов. При этом устанавливается соответствие между выходом нейронной сети и классом, который он представляет. Когда сети предъявляется некий образ, на одном из её выходов должен появиться признак того, что образ принадлежит этому классу. В то же время на других выходах должен быть признак того, что образ данному классу не принадлежит^[15]. Если на двух или более выходах есть признак принадлежности к классу, считается, что сеть «не уверена» в своём ответе.

Используемые архитектуры нейросетей

- Обучение с учителем:
 - Перцептрон
- Обучение без учителя:
 - Сети адаптивного резонанса
- Смешанное обучение:
 - Сеть радиально-базисных функций

Принятие решений и управление

Эта задача близка к задаче классификации. Классификации подлежат ситуации, характеристики которых поступают на вход нейронной сети. На выходе сети при этом должен появиться признак решения, которое она приняла. При этом в качестве входных сигналов используются различные критерии описания состояния управляемой системы^[16].

Используемые архитектуры нейросетей

- Обучение с учителем:
 - Перцептрон
- Смешанное обучение:
 - Сеть радиально-базисных функций

Кластеризация

Под кластеризацией понимается разбиение множества входных сигналов на классы, при том, что ни количество, ни признаки классов заранее не известны. После обучения такая сеть способна определять, к какому классу относится входной сигнал. Сеть также может сигнализировать о том, что входной сигнал не относится ни к одному из выделенных классов — это является признаком новых, отсутствующих в обучающей выборке, данных. Таким образом, подобная сеть *может выявлять новые, неизвестные ранее классы сигналов*. Соответствие между классами, выделенными сетью, и классами, существующими в предметной области, устанавливается человеком. Кластеризацию осуществляют, например, нейронные сети Кохонена.

Нейронные сети в простом варианте Кохонена не могут быть огромными, поэтому их делят на гиперслои (гиперколонки) и ядра (микроколонки). Если сравнивать с мозгом человека, то идеальное количество параллельных слоёв не должно быть более 112. Эти слои в свою очередь составляют

гиперслои (гиперколонку), в которой от 500 до 2000 микроколонок (ядер). При этом каждый слой делится на множество гиперколонок, пронизывающих насквозь эти слои. Микроколонки кодируются цифрами и единицами с получением результата на выходе. Если требуется, то лишние слои и нейроны удаляются или добавляются. Идеально для подбора числа нейронов и слоёв использовать суперкомпьютер. Такая система позволяет нейронным сетям быть пластичными.

Используемые архитектуры нейросетей

- Обучение без учителя:
 - Перцептрон
 - Самоорганизующаяся карта Кохонена
 - Нейронная сеть Кохонена
 - Сети адаптивного резонанса

Прогнозирование

Способности нейронной сети к прогнозированию напрямую следуют из её способности к обобщению и выделению скрытых зависимостей между входными и выходными данными. После обучения сеть способна предсказать будущее значение некой последовательности на основе нескольких предыдущих значений и (или) каких-то существующих в настоящий момент факторов. Следует отметить, что прогнозирование возможно только тогда, когда *предыдущие изменения действительно в какой-то степени определяют будущие*. Например, прогнозирование котировок акций на основе котировок за прошлую неделю может оказаться успешным (а может и не оказаться), тогда как прогнозирование результатов завтрашней лотереи на основе данных за последние 50 лет почти наверняка не даст никаких результатов.

Используемые архитектуры нейросетей

- Обучение с учителем:
 - Перцептрон
- Смешанное обучение:
 - Сеть радиально-базисных функций

Аппроксимация

Нейронные сети могут аппроксимировать непрерывные функции. Доказана обобщённая аппроксимационная теорема^[17]: с помощью линейных операций и каскадного соединения можно из произвольного нелинейного элемента получить устройство, вычисляющее любую непрерывную функцию с некоторой наперёд заданной точностью. Это означает, что нелинейная характеристика нейрона может быть произвольной: от сигмоидальной до произвольного волнового пакета или вейвлета, синуса или многочлена. От выбора нелинейной функции может зависеть сложность конкретной сети, но с любой нелинейностью сеть остаётся универсальным аппроксиматором и при правильном выборе структуры может достаточно точно аппроксимировать функционирование любого непрерывного автомата.

Используемые архитектуры нейросетей

- Обучение с учителем:
 - Перцептрон

- Смешанное обучение:
 - Сеть радиально-базисных функций

Сжатие данных и Ассоциативная память

Способность нейросетей к выявлению взаимосвязей между различными параметрами дает возможность выразить данные большой размерности более компактно, если данные тесно взаимосвязаны друг с другом. Обратный процесс — восстановление исходного набора данных из части информации — называется (авто)ассоциативной памятью. Ассоциативная память позволяет также восстанавливать исходный сигнал/образ из зашумленных/поврежденных входных данных. Решение задачи гетероассоциативной памяти позволяет реализовать память, адресуемую по содержанию^[16].

Используемые архитектуры нейросетей

- Обучение с учителем:
 - Перцептрон
- Обучение без учителя:
 - Нейронная сеть Хопфилда

Анализ данных

Используемые архитектуры нейросетей

- Обучение с учителем:
 - Перцептрон
- Обучение без учителя:
 - Перцептрон
 - Самоорганизующаяся карта Кохонена
 - Нейронная сеть Кохонена

Оптимизация

Используемые архитектуры нейросетей

- Обучение без учителя:
 - Самоорганизующаяся карта Кохонена
 - Нейронная сеть Кохонена

Этапы решения задач

- Сбор данных для обучения;
- Подготовка и нормализация данных;
- Выбор топологии сети;
- Экспериментальный подбор характеристик сети;
- Экспериментальный подбор параметров обучения;
- Собственно обучение;
- Проверка адекватности обучения;

- Корректировка параметров, окончательное обучение;
- Вербализация сети^[18] с целью дальнейшего использования.

Следует рассмотреть подробнее некоторые из этих этапов.

Сбор данных для обучения

Выбор данных для обучения сети и их обработка является самым сложным этапом решения задачи. Набор данных для обучения должен удовлетворять нескольким критериям:

- Репрезентативность — данные должны иллюстрировать истинное положение вещей в предметной области;
- Непротиворечивость — противоречивые данные в обучающей выборке приведут к плохому качеству обучения сети.

Исходные данные преобразуются к виду, в котором их можно подать на входы сети. Каждая запись в файле данных называется *обучающей парой* или *обучающим вектором*. Обучающий вектор содержит по одному значению на каждый вход сети и, в зависимости от типа обучения (с учителем или без), по одному значению для каждого выхода сети. Обучение сети на «сыром» наборе, как правило, не даёт качественных результатов. Существует ряд способов улучшить «восприятие» сети.

- *Нормировка* выполняется, когда на различные входы подаются данные разной размерности. Например, на первый вход сети подаются величины со значениями от нуля до единицы, а на второй — от ста до тысячи. При отсутствии нормировки значения на втором входе будут всегда оказывать существенно большее влияние на выход сети, чем значения на первом входе. При нормировке размерности всех входных и выходных данных сводятся воедино;
- *Квантование* выполняется над непрерывными величинами, для которых выделяется конечный набор дискретных значений. Например, квантование используют для задания частот звуковых сигналов при распознавании речи;
- *Фильтрация* выполняется для «зашумленных» данных.

Кроме того, большую роль играет само представление как входных, так и выходных данных. Предположим, сеть обучается распознаванию букв на изображениях и имеет один числовой выход — номер буквы в алфавите. В этом случае сеть получит ложное представление о том, что буквы с номерами 1 и 2 более похожи, чем буквы с номерами 1 и 3, что, в общем, неверно. Для того, чтобы избежать такой ситуации, используют топологию сети с большим числом выходов, когда каждый выход имеет свой смысл. Чем больше выходов в сети, тем большее расстояние между классами и тем сложнее их спутать.

Выбор топологии сети

Выбирать тип сети следует, исходя из постановки задачи и имеющихся данных для обучения. Для обучения с учителем требуется наличие для каждого элемента выборки «экспертной» оценки. Иногда получение такой оценки для большого массива данных просто невозможно. В этих случаях естественным выбором является сеть, обучающаяся без учителя (например, самоорганизующаяся карта Кохонена или нейронная сеть Хопфилда). При решении других задач (таких, как прогнозирование временных рядов) экспертная оценка уже содержится в исходных данных и может быть выделена при их обработке. В этом случае можно использовать многослойный перцептрон^[уточнить] или сеть Ворда.

Экспериментальный подбор характеристик с сети

После выбора общей структуры нужно экспериментально подобрать параметры сети. Для сетей, подобных перцептронной, это будет число слоев, число блоков в скрытых слоях (для сетей Ворда), наличие или отсутствие обходных соединений, передаточные функции нейронов. При выборе

количества слоев и нейронов в них следует исходить из того, что *способности сети к обобщению тем выше, чем больше суммарное число связей между нейронами*. С другой стороны, число связей ограничено сверху количеством записей в обучающих данных.

Экспериментальный подбор параметров обучения

После выбора конкретной топологии необходимо выбрать параметры обучения нейронной сети. Этот этап особенно важен для сетей, обучающихся с учителем. От правильного выбора параметров зависит не только то, насколько быстро ответы сети будут сходиться к правильным ответам. Например, выбор низкой скорости обучения увеличит время схождения, однако иногда позволяет избежать паралича сети. Увеличение момента обучения может привести как к увеличению, так и к уменьшению времени сходимости, в зависимости от формы поверхности ошибки. Исходя из такого противоречивого влияния параметров, можно сделать вывод, что их значения нужно выбирать экспериментально, руководствуясь при этом критерием завершения обучения (например, минимизация ошибки или ограничение по времени обучения).

Обучение сети

В процессе обучения сеть в определенном порядке просматривает обучающую выборку. Порядок просмотра может быть последовательным, случайным и т. д. Некоторые сети, обучающиеся без учителя (например, сети Хопфилда), просматривают выборку только один раз. Другие (например, сети Кохонена), а также сети, обучающиеся с учителем, просматривают выборку множество раз, при этом один полный проход по выборке называется *эпохой обучения*. При обучении с учителем набор исходных данных делят на две части — собственно обучающую выборку и тестовые данные; принцип разделения может быть произвольным. Обучающие данные подаются сети для обучения, а проверочные используются для расчета ошибки сети (проверочные данные никогда для обучения сети не применяются). Таким образом, если на проверочных данных ошибка уменьшается, то сеть действительно выполняет обобщение. Если ошибка на обучающих данных продолжает уменьшаться, а ошибка на тестовых данных увеличивается, значит, сеть перестала выполнять обобщение и просто «запоминает» обучающие данные. Это явление называется переобучением сети или оверфиттингом. В таких случаях обучение обычно прекращают. В процессе обучения могут проявиться другие проблемы, такие как паралич или попадание сети в локальный минимум поверхности ошибок. Невозможно заранее предсказать проявление той или иной проблемы, равно как и дать однозначные рекомендации к их разрешению.

Все выше сказанное относится только к итерационным алгоритмам поиска нейросетевых решений. Для них действительно нельзя ничего гарантировать и нельзя полностью автоматизировать обучение нейронных сетей. Однако, наряду с итерационными алгоритмами обучения, существуют не итерационные алгоритмы, обладающие очень высокой устойчивостью и позволяющие полностью автоматизировать процесс обучения.

Проверка адекватности обучения

Даже в случае успешного, на первый взгляд, обучения сеть не всегда обучается именно тому, чего от неё хотел создатель. Известен случай, когда сеть обучалась распознаванию изображений танков по фотографиям, однако позднее выяснилось, что все танки были сфотографированы на одном и том же фоне. В результате сеть «научилась» распознавать этот тип ландшафта, вместо того, чтобы «научиться» распознавать танки^[19]. Таким образом, сеть «понимает» не то, что от неё требовалось, а то, что проще всего обобщить.

Тестирование качества обучения нейросети необходимо проводить на примерах, которые не участвовали в её обучении. При этом число тестовых примеров должно быть тем больше, чем выше качество обучения. Если ошибки нейронной сети имеют вероятность близкую к одной миллиардной, то

и для подтверждения этой вероятности нужен миллиард тестовых примеров. Получается, что тестирование хорошо обученных нейронных сетей становится очень трудной задачей.

Классификация по типу входной информации

- Аналоговые нейронные сети (используют информацию в форме действительных чисел);
- Двоичные нейронные сети (оперируют с информацией, представленной в двоичном виде);
- Образные нейронные сети (оперируют с информацией, представленной в виде образов: знаков, иероглифов, символов).

Классификация по характеру обучения

- Обучение с учителем — выходное пространство решений нейронной сети известно;
- Обучение без учителя — нейронная сеть формирует выходное пространство решений только на основе входных воздействий. Такие сети называют самоорганизующимися;
- Обучение с подкреплением — система назначения штрафов и поощрений от среды.

Классификация по характеру настройки синапсов

- Сети с фиксированными связями (весовые коэффициенты нейронной сети выбираются сразу, исходя из условий задачи, при этом: $dW/dt = 0$, где W — весовые коэффициенты сети);
- Сети с динамическими связями (для них в процессе обучения происходит настройка синаптических связей, то есть $dW/dt \neq 0$, где W — весовые коэффициенты сети).

Классификация по времени передачи сигнала

В ряде нейронных сетей активирующая функция может зависеть не только от весовых коэффициентов связей w_{ij} , но и от времени передачи импульса (сигнала) по каналам связи τ_{ij} . Поэтому в общем виде активирующая (передающая) функция связи c_{ij} от элемента u_i к элементу u_j имеет вид:

$c_{ij}^* = f[w_{ij}(t), u_i^*(t - \tau_{ij})]$. Тогда *синхронной сетью* называют такую сеть, у которой время передачи τ_{ij} каждой связи равно либо нулю, либо фиксированной постоянной τ . *Асинхронной* называют такую сеть у которой время передачи τ_{ij} для каждой связи между элементами u_i и u_j своё, но тоже постоянное.

Классификация по характеру связей

Сети прямого распространения (Feedforward)

Все связи направлены строго от входных нейронов к выходным. Примерами таких сетей являются перцептрон Розенблатта, многослойный перцептрон, сети Ворда.

Рекуррентные нейронные сети

Сигнал с выходных нейронов или нейронов скрытого слоя частично передается обратно на входы нейронов входного слоя (обратная связь). Рекуррентная сеть Хопфилда «фильтрует» входные данные, возвращаясь к устойчивому состоянию и, таким образом, позволяет решать задачи компрессии данных и построения ассоциативной памяти^[20]. Частным случаем рекуррентных сетей являются двунаправленные сети. В таких сетях между слоями существуют связи как в направлении от входного слоя к выходному, так и в обратном. Классическим примером является Нейронная сеть Коско.

Радиально-базисные функции

Разработаны искусственные нейронные сети, использующие в качестве активационных функций радиально-базисные (также называются RBF-сетями). Общий вид радиально-базисной функции:

$$f(\mathbf{x}) = \phi\left(\frac{\mathbf{x}^2}{\sigma^2}\right), \text{ например, } f(\mathbf{x}) = e^{-\frac{\mathbf{x}^2}{\sigma^2}},$$

где \mathbf{x} — вектор входных сигналов нейрона, σ — ширина окна функции, $\phi(y)$ — убывающая функция (чаще всего, равная нулю вне некоторого отрезка).

Радиально-базисная сеть характеризуется тремя особенностями:

1. Единственный скрытый слой;
2. Только нейроны скрытого слоя имеют нелинейную активационную функцию;
3. Синаптические веса связей входного и скрытого слоев равны единице.

Самоорганизующиеся карты

Такие сети представляют собой соревновательную нейронную сеть с обучением без учителя, выполняющую задачу визуализации и кластеризации. Является методом проецирования многомерного пространства в пространство с более низкой размерностью (чаще всего, двумерное), применяется также для решения задач моделирования, прогнозирования и др. Является одной из версий нейронных сетей Кохонена^[21]. Самоорганизующиеся карты Кохонена служат, в первую очередь, для визуализации и первоначального («разведывательного») анализа данных^[22].

Сигнал в сеть Кохонена поступает сразу на все нейроны, веса соответствующих синапсов интерпретируются как координаты положения узла, и выходной сигнал формируется по принципу «победитель забирает всё» — то есть ненулевой выходной сигнал имеет нейрон, ближайший (в смысле весов синапсов) к подаваемому на вход объекту. В процессе обучения веса синапсов настраиваются таким образом, чтобы узлы решетки «располагались» в местах локальных сгущений данных, то есть описывали кластерную структуру облака данных, с другой стороны, связи между нейронами соответствуют отношениям соседства между соответствующими кластерами в пространстве признаков.

Удобно рассматривать такие карты как двумерные сетки узлов, размещенных в многомерном пространстве. Изначально самоорганизующаяся карта представляет собой сетку из узлов, соединенную между собой связями. Кохонен рассматривал два варианта соединения узлов — в прямоугольную и гексагональную сетку — отличие состоит в том, что в прямоугольной сетке каждый узел соединён с 4-мя соседними, а в гексагональной — с шестью ближайшими узлами. Для двух таких сеток процесс построения сети Кохонена отличается лишь в том месте, где перебираются ближайшие к данному узлу соседи.

Начальное вложение сетки в пространство данных выбирается произвольным образом. В авторском пакете SOM_PAK предлагаются варианты случайного начального расположения узлов в пространстве и вариант расположения узлов в плоскости. После этого узлы начинают перемещаться в пространстве согласно следующему алгоритму:

1. Случайным образом выбирается точка данных \mathbf{x} .
2. Определяется ближайший к \mathbf{x} узел карты (BMU — Best Matching Unit).
3. Этот узел перемещается на заданный шаг по направлению к \mathbf{x} . Однако он перемещается не один, а увлекает за собой определённое количество ближайших узлов из некоторой окрестности на карте. Из всехдвигающихся узлов наиболее сильно смещается центральный — ближайший к точке данных — узел, а остальные испытывают тем меньшие смещения, чем дальше они от BMU. В настройке карты различают два этапа — этап грубой (*ordering*) и этап тонкой (*fine-tuning*) настройки. На первом этапе выбираются большие значения окрестностей и движение узлов носит коллективный характер — в результате карта «расправляется» и грубым образом отражает структуру данных; на этапе тонкой настройки радиус окрестности равен 1-2 и настраиваются уже

индивидуальные положения узлов. Кроме этого, величина смещения равномерно затухает со временем, то есть она велика в начале каждого из этапов обучения и близка к нулю в конце.

4. Алгоритм повторяется определенное число эпох (понятно, что число шагов может сильно изменяться в зависимости от задачи).

Известные типы сетей

- Перцептрон Розенблатта;
- Сплайн-модель Хакимова;
- Многослойный перцептрон Розенблатта;
- Многослойный перцептрон Румельхарта;
- Сеть Джордана;
- Сеть Элмана;
- Сеть Хэмминга;
- Сеть Ворда;
- Сеть Хопфилда;
- Сеть Кохонена;
- Нейронный газ^[23];
- Когнитрон;
- Неокогнитрон;
- Хаотическая нейронная сеть;
- Осцилляторная нейронная сеть;
- Сеть встречного распространения;
- Сеть радиально-базисных функций (RBF-сеть);
- Сеть обобщенной регрессии;
- Сеть Д.Смирнова;
- Вероятностная сеть;
- Вероятностная нейронная сеть Решетова;
- Сиамская нейронная сеть;
- Сети адаптивного резонанса;
- Свёрточная нейронная сеть (англ. *convolutional neural network*);
- Нечеткий многослойный перцептрон;
- Импульсная нейронная сеть.

Отличия от машин с архитектурой фон Неймана

Вычислительные системы, основанные на искусственных нейронных сетях, обладают рядом качеств, которые отсутствуют в машинах с архитектурой фон Неймана (но присущи мозгу человека):

- Массовый параллелизм;
- Распределённое представление информации и вычисления;
- Способность к обучению и обобщению;
- Адаптивность;
- Свойство контекстуальной обработки информации;
- Толерантность к ошибкам;
- Низкое энергопотребление.

Примеры приложений

Предсказание финансовых временных рядов

Входные данные — курс акций за год. Задача — определить завтрашний курс. Проводится следующее преобразование — выстраивается в ряд курс за сегодня, вчера, за позавчера. Следующий ряд — смещается по дате на один день и так далее. На полученном наборе обучается сеть с 3 входами и одним выходом — то есть выход: курс на дату, входы: курс на дату минус 1 день, минус 2 дня, минус 3 дня. Обученной сети подаем на вход курс за сегодня, вчера, позавчера и получаем ответ на завтра. Нетрудно заметить, что в этом случае сеть просто выведет зависимость одного параметра от трёх предыдущих. Если желательно учитывать ещё какой-то параметр (например, общий индекс по отрасли), то его надо добавить как вход (и включить в примеры), переобучить сеть и получить новые результаты. Для наиболее точного обучения стоит использовать метод ОРО, как наиболее предсказуемый и несложный в реализации.

Психодиагностика

Серия работ М. Г. Доррера с соавторами посвящена исследованию вопроса о возможности развития психологической интуиции у нейросетевых экспертных систем^{[24][25]}. Полученные результаты дают подход к раскрытию механизма интуиции нейронных сетей, проявляющейся при решении ими психодиагностических задач. Создан нестандартный для компьютерных методик *интуитивный* подход к психодиагностике, заключающийся в исключении построения *описанной реальности*. Он позволяет сократить и упростить работу над психодиагностическими методиками.

Хемоинформатика

Нейронные сети широко используются в химических и биохимических исследованиях^[26]. В настоящее время нейронные сети являются одним из самых распространенных методов хемоинформатики для поиска количественных соотношений структура-свойство^{[27][28]}, благодаря чему они активно используются как для прогнозирования физико-химических свойств и биологической активности химических соединений, так и для направленного дизайна химических соединений и материалов с заранее заданными свойствами, в том числе при разработке новых лекарственных препаратов.

Нейроуправление

Нейронные сети успешно применяются для синтеза систем управления динамическими объектами^{[29][30]}. Нейросети обладают рядом уникальных свойств, которые делают их мощным инструментом для создания систем управления: способностью к обучению на примерах и обобщению данных, способностью адаптироваться к изменению свойств объекта управления и внешней среды, пригодностью для синтеза нелинейных регуляторов, высокой устойчивостью к повреждениям своих элементов в силу изначально заложенного в нейросетевую архитектуру параллелизма.

Экономика

Алгоритмы искусственных нейронных сетей нашли широкое применение в экономике^[31]. С помощью нейронных сетей решается задача разработки алгоритмов нахождения аналитического описания закономерностей функционирования экономических объектов (предприятие, отрасль, регион). Эти алгоритмы применяются к прогнозированию некоторых «выходных» показателей объектов. Применение нейросетевых методов позволяет решить некоторые проблемы экономико-статистического моделирования, повысить адекватность математических моделей, приблизить их к экономической реальности^[32]. Поскольку экономические, финансовые и социальные системы очень сложны и являются результатом человеческих действий и противодействий, создание полной математической модели с учётом всех возможных действий и противодействий является очень сложной (если разрешимой) задачей. В системах подобной сложности естественным и наиболее эффективным является использование моделей, которые напрямую имитируют поведение общества и экономики. Именно это способна предложить методология нейронных сетей^[33].

См. также

- Оптические нейронные сети
- Искусственный интеллект
- Нейрокомпьютер
- Blue Brain Project
- Модель биологического нейрона
- Когнитивистика

Примечания

1. Мак-Каллок У. С., Питтс В. Логическое исчисление идей, относящихся к нервной

активности (<http://neuro.net.ua/pub/mcculloch.html>)/ Автоматы / Под ред. К. Э. Шеннона и Дж.

- Маккарти. — М.: Изд-во иностр. лит., 1956. — С. 363—384. (Перевод английской статьи 1943 г)
2. ↑ Горбань А. Н. Кто мы, куда мы идём, как путь наш измерить? (http://www.library.mephi.ru/data/scientific-sessions/2000/Book_Diss/33.html) Пленарный доклад на открытии конференции *Нейроинформатика-99* (МИФИ, 20 января 1999). Журнальный вариант: Горбань А. Н. Нейроинформатика: кто мы, куда мы идём, как путь наш измерить // Вычислительные технологии. — М.: Машиностроение. — 2000. — № 4. — С. 10—14. = Gorban A.N. Neuroinformatics: What are we, where are we going, how to measure our way? (<http://arxiv.org/abs/cond-mat/0307346>) The Lecture at the USA-NIS Neurocomputing Opportunities Workshop, Washington DC, July 1999 (Associated with IJCNN'99).
 3. ↑ Н. Винер. Кибернетика. 2-е изд., 1961, т. I.
 4. ↑ Голубев, 2007, с. 4.
 5. ↑ Pattern Recognition and Adaptive Control. BERNARD WIDROW (<http://www-isl.stanford.edu/~widrow/papers/j1964patternrecognition.pdf>)
 6. ↑ Уидроу Б., Стирнс С. Адаптивная обработка сигналов. — М.: Радио и связь, 1989. — 440 с.
 7. ↑ Петров А. П. О возможностях перцептрона // Известия АН СССР, Техническая кибернетика. — 1964. — № 6.
 8. ↑ Бонгард М. М. Проблемы узнавания. — М.: Физматгиз, 1967.
 9. ↑ Голубев, 2007, с. 5.
 10. ↑ Хакимов Б. Б. Моделирование корреляционных зависимостей сплайнами на примерах в экологии и экологии. — М.: Изд-во Моск. ун-та; СПб.: Нева, 2003. — 144 с.
 11. ↑ Werbos P. J. Beyond regression: New tools for prediction and analysis in the behavioral sciences. — Ph. D. thesis, Harvard University Cambridge, MA, 1974.
 12. ↑ Галушкин А. И. Синтез многослойных систем распознавания образов. — М.: Энергия, 1974.
 13. ↑ Rumelhart D. E., Hinton G. E., Williams R. J. Learning Internal Representations by Error Propagation // Parallel Distributed Processing. Vol. 1. — Cambridge, MA: MIT Press, 1986. P318—362.
 14. ↑ Барцев С. И., Ожнин В. А. Адаптивные сети обработки информации. — Красноярск: Ин-т физики СО АН СССР 1986. Препринт N 5Б. — 20 с.
 15. ↑ Такой вид кодирования иногда называют кодом «1 из N»
 16. Открытые системы — введение в нейросети (<http://www.osp.ru/os/1997/04/16.htm>)
 17. ↑ Горбань А. Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей (<http://neuroschool.narod.ru/pub/sibzhvmt98.pdf>) // Сибирский журнал вычислительной математики, 1998, т. 1, № 1. — С. 12—24.
 18. ↑ Миркес Е. М. Логически прозрачные нейронные сети и производство явных знаний из данных (<http://www.intuit.ru/department/expert/neuroinf/9/>) Нейроинформатика / А. Н. Горбань, В. Л. Дунин-Барковский, А. Н. Кирдин и др. — Новосибирск: Наука. Сибирское предприятие РАН, 1998. — 296 с. — ISBN 5-02-031410-2
 19. ↑ Упоминание этой истории в журнале «Популярная механика» (<http://www.popmech.ru/article/2233-po-obrazu-i-podobiyu/>)
 20. ↑ INTUIT.ru — Рекуррентные сети как ассоциативные запоминающие устройства (<http://www.intuit.ru/department/expert/neuro/10/>)
 21. ↑ Kohonen T. Self-Organizing Maps. 3rd edition. — Berlin — New York: Springer-Verlag, 2001/ — ISBN 0-387-51387-6 ISBN 3-540-67921-9
 22. ↑ Зиновьев А. Ю. Визуализация многомерных данных (<http://pca.narod.ru/ZINANN.htm>) — Красноярск: Изд-во Красноярского гос. техн. ун-та, 2000. — 180 с.
 23. ↑ Martinetz T. M., Berkovich S. G., Schulten K. J. Neural-gas network for vector quantization and its application to time-series prediction (<http://pca.narod.ru/MartinesShultenNeuralGas1993.pdf>) // IEEE Trans. on Neural Networks, 1993, No. 4. — P558—569. На сайте PCA (<http://pca.narod.ru/>)
 24. ↑ Gorban A. N., Rossiyev D. A., Dorer M. G. MultiNeuron — Neural Networks Simulator For Medical, Physiological, and Psychological Applications (<http://arxiv.org/ftp/q-bio/papers/0411/0411034.pdf>) // Wcnn'95, Washington, D.C.: World Congress on Neural Networks, 1995. International Neural Network Society Annual Meeting : Renaissance Hotel, Washington, D.C., USA, July 17-21, 1995.
 25. ↑ Доррер М. Г. Психологическая интуиция искусственных нейронных сетей (<http://psyfactor.org/lib/dorrer-0.htm>), Дисс. ... 1998. Другие копии онлайн: [1] (<http://www.tnu.in.ua/study/books.php?do=file&id=1501>) [2] (<http://lib.sibnet.ru/referat/6923>)
 26. ↑ Баскин И. И., Палюин В. А., Зефиоров Н. С. Применение искусственных нейронных сетей в химических и биохимических исследованиях (<http://www.chem.msu.su/rus/vmgu/995/323.pdf>) // Вестн. Моск. ун-та. Сер. 2. Химия 1999. Т. 40. № 5.
 27. ↑ Гальберштам Н. М., Баскин И. И., Палюин В. А., Зефиоров Н. С. Нейронные сети как метод поиска зависимостей структура — свойство органических соединений // Успехи химии. — 2003. — Т. 72, № 7. — С. 706-727.
 28. ↑ Баскин И. И., Палюин В. А., Зефиоров Н. С. Многослойные перцептроны в исследовании зависимостей «структура-свойство» для органических соединений // Русский химический журнал (Журнал Российского химического общества им. Д. И. Менделеева). — 2006. — Т. 50. — С. 86—96.
 29. ↑ Сигеру, Марзуки, Рубия, 2000.

30. ↑ Чернодуб А. Н., Дзюба Д. А. Обзор методов нейрорууправления (http://ailen.org/wp-content/uploads/2011/08/2011_NeuroControl_Survey.pdf) // *Проблемы программирования*. — 2011. — No 2. — С. 79—94.
31. ↑ Калацкая Л. В., Новиков В. А., Садов В. С. Организация и обучение искусственных нейронных сетей: Экспериментальное учеб. пособие. — Минск: Изд-во БГУ, 2003. — 72 с.
32. ↑ Кенин А. М., Мазуров В. Д. Опыт применения нейронных сетей в экономических задачах (<http://www.uralstars.com/Docs/Editor/Neuro.htm>)
33. ↑ [3] (<http://www.intuit.ru/department/itmngt/iteconomy/11/2.html>) Нейронные сети в экономике

Литература

- Беркинблит М. Б. Нейронные сети. — М.: МИРОС и ВЗМШ РАО, 1993. — 96 с. — ISBN 5-7084-0026-9.
- Вороновский Г. К., Махотило К. В., Петрашев С. Н., Сергеев С. А. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности. — Харьков: Основа, 1997. — 112 с. — ISBN 5-7768-0293-8.
- Голубев Ю. Ф. Нейросетевые методы в мехатронике. — М.: Изд-во Моск. унта, 2007. — 157 с. — ISBN 978-5-211-05434-9.
- Горбань А.Н. Обучение нейронных сетей. — М.: СССР-США СП «Параграф», 1990. — 160 с.
- Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. — Новосибирск: Наука, 1996. — 276 с. — ISBN 5-02-031196-0.
- Горбань А. Н., Дунин-Барковский В. Л. и др. Нейроинформатика. — Новосибирск: Наука, 1998.
- Еремин Д. М., Гарцеев И. Б. Искусственные нейронные сети в интеллектуальных системах управления. — М.: МИРЭА, 2004. — 75 с. — ISBN 5-7339-0423-2.
- Каллан Р. Основные концепции нейронных сетей = The Essence of Neural Networks First Edition. — М.: Вильямс, 2001. — 288 с. — ISBN 5-8459-0210-X.
- Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. — М.: Горячая линия - Телеком, 2001. — 382 с. — ISBN 5-93517-031-0.
- Миркес Е. М. Нейрокомпьютер. Проект стандарта. — Новосибирск: Наука, 1999. — 337 с. — ISBN 5-02-031409-9. Другие копии онлайн: Нейрокомпьютер. Проект стандарта.
- Осовский Станислав. Нейронные сети для обработки информации = Sieci neuronowe do przetwarzania informacji (польск.) / Перевод И. Д. Рудинского. — М.: Финансы и статистика, 2004. — 344 с. — ISBN 5-279-02567-4.
- Савельев А. В. На пути к общей теории нейросетей. К вопросу о сложности // Нейрокомпьютеры: разработка, применение. — 2006. — № 4—5. — С. 4—14.
- Сигеру Омату, Марзуки Халид, Рубия Юсоф. Нейрорууправление и его приложения = Neuro-Control and its Applications. 2-е изд. — М.: ИПРЖР, 2000. — 272 с. — ISBN 5-93108-006-6.
- Тадеусевич Рышард, Боровик Барбара, Гончаж Томаш, Леппер Бартош. Элементарное введение в технологию нейронных сетей с примерами программ / Перевод И. Д. Рудинского. — М.: Горячая линия — Телеком, 2011. — 408 с. — ISBN 978-5-9912-0163-6..
- Терехов В. А., Ефимов Д. В., Тюкин И. Ю. Нейросетевые системы управления. — М.: Высшая школа, 2002. — 184 с. — ISBN 5-06-004094-1.
- Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика = Neural Computing. Theory and Practice. — М.: Мир, 1992. — 240 с. — ISBN 5-03-002115-9.
- Хайкин С. Нейронные сети: полный курс = Neural Networks: A Comprehensive Foundation. 2-е изд. — М.: Вильямс, 2006. — 1104 с. — ISBN 0-13-273350-1.
- Ясницкий Л. Н. Введение в искусственный интеллект. — М.: Издат. центр «Академия», 2005. — 176 с. — ISBN 5-7695-1958-4.

Ссылки

- Neural Networks в каталоге ссылок Open Directory Project (dmoz)
- Учебник по искусственным нейронным сетям
- Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning (англ.). MIT Press (2016). — Книга по машинному обучению и, в частности, по глубокому обучению.

Источник — «https://ru.wikipedia.org/w/index.php?title=Искусственная_нейронная_сеть&oldid=86677002»

- Эта страница последний раз была отредактирована 22 июля 2017 в 20:22.
- Текст доступен по лицензии Creative Commons Attribution-ShareAlike; в отдельных случаях могут действовать дополнительные условия.
Wikipedia® — зарегистрированный товарный знак некоммерческой организации Wikimedia Foundation, Inc.