

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

Нейрокомпьютеринг.

Лекция 14. Нейрокомпьютеринг. Введение

О.Г. Чанышев

Содержание

1 Основные понятия	3
1.1 Модель технического нейрона	3
1.2 Архитектура нейронной сети	3
2 Обучение нейросетей	3
2.1 Парадигмы обучения	4
2.2 Свойства ИНС, связанные с обучением	4
2.3 Основные правила обучения	4
2.3.1 Правило коррекции по ошибке	4
2.3.2 Обучение Больцмана	4
2.3.3 Правило Хебба	5
2.3.4 Обучение методом соревнования	5

Введение

Как уже говорилось во вступительной лекции, с самого начала исследователи Мозга пошли по двум основным направлениям: одни пытались воспроизвести результаты деятельности Мозга, безотносительно его конструкции (направление "черного ящика"), другие пытались смоделировать конструкцию Мозга. Последнее направление, стартовавшее с создания модели искусственного нейрона, привело к настоящему времени к множеству различных структур искусственных нейронных сетей (ИНС), реализованных как в "железе", так и программно. В первых же работах по моделированию конструкции Мозга на ЭВМ выяснилось, что эти модели не только воспроизводят некоторые функции мозга, но имеют собственную ценность.

Некоторые ученые считают ИНС членом парадигмы "Искусственный интеллект", другие - отдельным направлением исследований. Так или иначе, на вопрос о том, какие знания необходимы для глубокого изучения проблематики ИНС, можно ответить: те же, что и для глубокого понимания проблемы ИИ, а именно:

гносеологии, нейрофизиологии, психологии, физики (статистической механики), теории управления, теории вычислений, проблем искусственного интеллекта, статистики/математики, распознавания образов, компьютерного зрения, параллельных вычислений и аппаратных средств.

Этот симбиоз жизненно необходим для исследований по нейронным сетям. Исследования ИНС, в свою очередь, способствуют прогрессу в познании перечисленных предметных областей.

Возможно, самым полным обобщением сказанного является определение ИИ, данное В.М. Сергеевым [1]: «искусственный интеллект - это экспериментальная философия, фантастичекая еще недавно возможность проверки самых тонких и абстрактных гипотез о природе человеческого разума».

Перечислим некоторые из проблем, решаемые в контексте ИНС.

- Классификация и кластеризация (классификация «без учителя»).
- Аппроксимация функций. Имеется обучающая выборка

$$((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$$

(пары данных вход-выход), которая генерируется неизвестной функцией (x) , искаженной шумом. Задача состоит в нахождении оценки неизвестной функции (x) .

- Прогноз (предсказание). Заданы n дискретных отсчетов

$$y(t_1), y(t_2), \dots, y(t_n)$$

в последовательные моменты времени t_1, t_2, \dots, t_n . Задача состоит в предсказании значения $y(t_{n+1})$ в некоторый будущий момент времени t_{n+1} .

- Задачей алгоритма оптимизации является нахождение такого решения, которое удовлетворяет системе ограничений и максимизирует или минимизирует целевую функцию. Задача коммивояжера является классическим примером задачи оптимизации.
- Ассоциативная память, или память, адресуемая по содержанию, доступна по указанию заданного содержания. Содержимое памяти может быть вызвано даже по частичному входу или искаженному содержанию. Ассоциативная память чрезвычайно желательна при создании мультимедийных информационных баз данных.
- Управление. Динамическая система заданна совокупностью

$$(u(t), y(t)),$$

где $u(t)$ является входным управляющим воздействием, а $y(t)$ - выходом системы в момент времени t . В системах управления с эталонной моделью целью управления является расчет такого входного воздействия $u(t)$, при котором система следует по желаемой траектории, диктуемой эталонной моделью. Примером является оптимальное управление двигателем.

Краткий исторический обзор

Исследования в области ИНС пережили три периода активизации. Первый пик в 40-х годах обусловлен пионерской работой МакКаллока и Питтса [4]. Второй возник в 60-х благодаря теореме сходимости перцептрона Розенблатта [5] и работе Минского и Пейперта [6], указавшей ограниченные возможности простейшего перцептрона. Результаты Минского и Пейперта погасили энтузиазм большинства исследователей, особенно тех, кто работал в области вычислительных наук. Возникшее в исследованиях по нейронным сетям затишье продлилось почти 20 лет. С начала 80-х годов ИНС вновь привлекли интерес исследователей, что связано с энергетическим подходом Хопфилда [7] и алгоритмом обратного распространения для обучения

Рис. 1: Искусственный нейрон

Рис. 2: Грубая систематизация архитектур сетей

многослойного перцептрона (многослойные сети прямого распространения), впервые предложенного Вербосом [8] и независимо разработанного рядом других авторов. Алгоритм получил известность благодаря Румельхарту [9] в 1986 году Андерсон и Розенфельд [10] подготовили подробную историческую справку о развитии ИНС.

Наращение научной активности носило взрывной характер. В 1987 г. было проведено четыре крупных совещания по искусственным нейронным сетям и опубликовано свыше 500 научных сообщений. Таким образом, еще и еще раз подтвердился закон Кларка (писатель и ученый Артур Кларк), гласящий: *если крупный уважаемый ученый говорит, что нечто может быть выполнено, то он почти всегда прав. Если же ученый говорит, что это не может быть выполнено, то он почти всегда не прав.*

1 Основные понятия

1.1 Модель технического нейрона

МакКаллок и Питтс [4] предложили использовать бинарный пороговый элемент в качестве модели искусственного нейрона (рис. 1) На вход искусственного нейрона поступает некоторое множество сигналов x_1, x_2, \dots, x_n , каждый из которых является выходом другого нейрона. Каждый вход умножается на соответствующий вес (w_1, w_2, \dots, w_n) и поступает на суммирующий блок, обозначенный S . Каждый вес соответствует "силе" одной биологической синаптической связи. (Множество весов в совокупности обозначается вектором W .) Суммирующий блок складывает взвешенные входы алгебраически, создавая выход NET . В векторном представлении: $NET = XW$. Сигнал NET далее, как правило, преобразуется активационной функцией F и дает выходной нейронный сигнал OUT .

Активационная функция может быть обычной линейной функцией $OUT = K(NET)$, где K - постоянная, пороговой функцией $OUT = 1$, если $NET > T$, $OUT = 0$ в остальных случаях, где T - некоторая постоянная пороговая величина, или же функцией, более точно моделирующей нелинейную передаточную характеристику биологического нейрона и представляющей нейронной сети большие возможности. Если блок F сужает диапазон изменения величины NET так, что при любых значениях NET значения OUT принадлежат некоторому конечному интервалу, то F называется "сжимающей" функцией. В качестве "сжимающей" функции часто используется логистическая или "сигмоидальная" (S -образная) функция. Эта функция математически выражается как $F(x) = 1/(1 + e^{-x})$. Таким образом,

1.2 Архитектура нейронной сети

ИНС может рассматриваться как направленный граф со взвешенными связями, в котором искусственные нейроны являются узлами. По архитектуре связей ИНС могут быть сгруппированы в два класса (рис. 2): сети прямого распространения, в которых графы не имеют петель, и рекуррентные сети, или сети с обратными связями.

2 Обучение нейросетей

Способность к обучению является фундаментальным свойством мозга. В контексте ИНС процесс обучения может рассматриваться как настройка архитектуры сети и весов связей для эффективного выполнения специальной задачи. Обычно нейронная сеть должна настроить веса связей по имеющейся обучающей выборке. Функционирование сети улучшается по мере итеративной настройки весовых коэффициентов. Свойство сети обучаться на примерах делает их более привлекательными по сравнению с системами, которые следуют определенной системе правил функционирования, сформулированной экспертами.

Для конструирования процесса обучения, прежде всего, необходимо иметь модель внешней среды, в которой функционирует нейронная сеть - знать доступную для сети информацию. Эта модель определяет парадигму обучения. Во-вторых, необходимо понять, как модифицировать весовые параметры сети - какие правила обучения управляют процессом настройки. Алгоритм обучения означает процедуру, в которой используются правила обучения для настройки весов.

2.1 Парадигмы обучения

Существуют три парадигмы обучения: «с учителем», «без учителя» (самообучение) и смешанная. В первом случае нейронная сеть располагает правильными ответами (выходами сети) на каждый входной пример. Веса настраиваются так, чтобы сеть производила ответы как можно более близкие к известным правильным ответам. Усиленный вариант обучения с учителем предполагает, что известна только критическая оценка правильности выхода нейронной сети, но не сами правильные значения выхода. Обучение без учителя не требует знания правильных ответов на каждый пример обучающей выборки. В этом случае раскрывается внутренняя структура данных или корреляции между образцами в системе данных, что позволяет распределить образцы по категориям. При смешанном обучении часть весов определяется посредством обучения с учителем, в то время как остальная получается с помощью самообучения.

2.2 Свойства ИНС, связанные с обучением

Теория обучения рассматривает три фундаментальных свойства, связанных с обучением по примерам: емкость, сложность образцов и вычислительная сложность. Под емкостью понимается, сколько образцов может запомнить сеть, и какие функции и границы принятия решений могут быть на ней сформированы. Сложность образцов определяет число обучающих примеров, необходимых для достижения способности сети к обобщению. Слишком малое число примеров может вызвать "переобученность" сети, когда она хорошо функционирует на примерах обучающей выборки, но плохо - на тестовых примерах, подчиненных тому же статистическому распределению. Известны 4 основных типа правил обучения: коррекция по ошибке, машина Больцмана, правило Хебба и обучение методом соревнования.

2.3 Основные правила обучения

2.3.1 Правило коррекции по ошибке

При обучении с учителем для каждого входного примера задан желаемый выход d . Реальный выход сети y может не совпадать с желаемым. Принцип коррекции по ошибке при обучении

Рис. 3: Пример обучения методом соревнования:(А) перед обучением; (Б) после обучения

состоит в использовании сигнала $(d - y)$ для модификации весов, обеспечивающей постепенное уменьшение ошибки. Обучение имеет место только в случае, когда перцептрон ошибается.

2.3.2 Обучение Больцмана

Этот тип обучения представляет собой стохастическое правило обучения, которое следует из информационных теоретических и термодинамических принципов. Целью обучения Больцмана является такая настройка весовых коэффициентов, при которой состояния видимых нейронов удовлетворяют желаемому распределению вероятностей. Обучение Больцмана может рассматриваться как специальный случай коррекции по ошибке, в котором под ошибкой понимается расхождение корреляций состояний в двух режимах.

2.3.3 Правило Хебба

Самым старым обучающим правилом является постулат обучения Хебба. Хебб опирался на следующие нейрофизиологические наблюдения: если нейроны с обеих сторон синапса активизируются одновременно и регулярно, то сила синаптической связи возрастает. Важной особенностью этого правила является то, что изменение синаптического веса зависит только от активности нейронов, которые связаны данным синапсом.

В ИНС, использующей обучение по Хеббу, наращивание весов определяется произведением уровней возбуждения передающего и принимающего нейронов. Это можно записать как

$$w_{ij}(n + 1) = w(n) + \alpha \times OUT_i \times OUT_j,$$

где $w_{ij}(n)$ - значение веса от нейрона i к нейрону j до подстройки, $w_{ij}(n + 1)$ - значение веса от нейрона i к нейрону j после подстройки, α - коэффициент скорости обучения, OUT_i - выход нейрона i и вход нейрона j , OUT_j - выход нейрона j .

2.3.4 Обучение методом соревнования

В отличие от обучения Хебба, в котором множество выходных нейронов могут возбуждаться одновременно, при соревновательном обучении выходные нейроны соревнуются между собой за активизацию. Это явление известно как правило "победитель берет все". Подобное обучение имеет место в биологических нейронных сетях. Обучение посредством соревнования позволяет кластеризовать входные данные: подобные примеры группируются сетью в соответствии с корреляциями и представляются одним элементом.

При обучении модифицируются только веса "победившего" нейрона. Эффект этого правила достигается за счет такого изменения сохраненного в сети образца (вектора весов связей победившего нейрона), при котором он становится чуть ближе к входному примеру. На рис. 3 дана геометрическая иллюстрация обучения методом соревнования. Входные векторы нормализованы и представлены точками на поверхности сферы. Векторы весов для трех нейронов инициализированы случайными значениями. Их начальные и конечные значения после обучения отмечены на рис. 3.a и 3.b соответственно. Каждая из трех групп примеров обнаружена одним из выходных нейронов, чей весовой вектор настроился на центр тяжести обнаруженной группы.

Можно заметить, что сеть никогда не перестанет обучаться, если параметр скорости обучения не равен 0. Некоторый входной образец может активизировать другой выходной нейрон на последующих итерациях в процессе обучения. Это ставит вопрос об устойчивости обучающей системы. *Система считается устойчивой, если ни один из примеров обучающей выборки не изменяет своей принадлежности к категории после конечного числа итераций обучающего процесса.* Один из способов достижения стабильности состоит в постепенном уменьшении до 0 параметра скорости обучения. Однако это искусственное торможение обучения вызывает другую проблему, называемую пластичностью и связанную со способностью к адаптации к новым данным. Эти особенности обучения методом соревнования известны под названием дилеммы стабильности-пластичности Гроссберга.

Следующая лекция

Лекция 15. Персептроны