

Содержание:

image not found or type unknown



Введение

В последние десятилетия в мире бурно развивается новая прикладная область математики, специализирующаяся на искусственных нейронных сетях. Актуальность исследований в этом направлении подтверждается массой различных применений нейросетей. Это автоматизация процессов распознавания образов, адаптивное управление, аппроксимация функционалов, прогнозирование, создание экспертных систем, организация ассоциативной памяти и многие другие приложения. С помощью нейросетей можно, например, предсказывать показатели биржевого рынка, выполнять распознавание оптических или звуковых сигналов, создавать самообучающиеся системы, способные управлять автомобилем при парковке или синтезировать речь по тексту. В то время как на западе применение ИС уже достаточно обширно, у нас это еще в некоторой степени экзотика – российские фирмы, использующие ИС в практических целях, наперечет.

Широкий круг задач, решаемый нейронными сетями, не позволяет в настоящее время создавать универсальные, мощные сети, вынуждая разрабатывать специализированные сети, функционирующие по различным алгоритмам. Тем не менее, тенденции развития нейросетей растут с каждым годом.

Цель моей работы – разбор базовых понятий, связанных с изучением нейронных сетей, а также выявление перспектив развития.

1. Базовые понятия искусственного нейрона

1.1 Структура искусственного нейрона

Искусственный нейрон имитирует в первом приближении свойства биологического нейрона. На вход искусственного нейрона поступает некоторое множество

сигналов, каждый из которых является выходом другого нейрона. Каждый вход умножается на соответствующий вес, аналогичный синаптической силе, и все произведения суммируются, определяя уровень активации нейрона. На рис.1 представлена модель, реализующая эту идею.

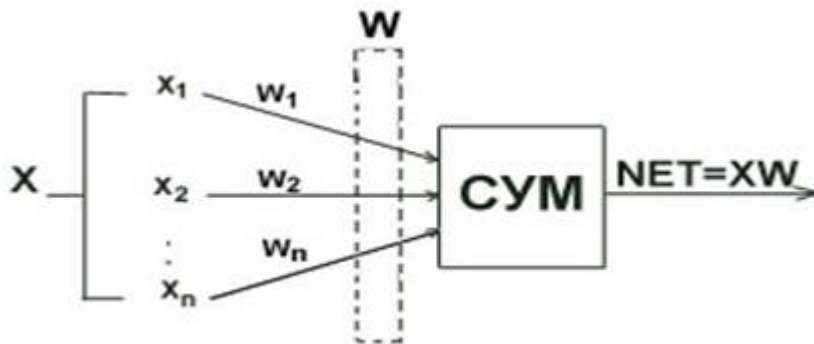


рис.1.Искусственный нейрон

Хотя сетевые парадигмы весьма разнообразны, в основе почти всех их лежит эта конфигурация. Здесь множество входных сигналов, обозначенных $x_1, x_2, x_3 \dots x_n$, поступает на искусственный нейрон. Эти входные сигналы, в совокупности обозначаемые вектором X , соответствуют сигналам, приходящим в синапсы биологического нейрона. Каждый сигнал умножается на соответствующий вес $w_1, w_2, w_3 \dots w_n$, и поступает на суммирующий блок, обозначенный СУМ. Каждый вес соответствует "силе" одной биологической синаптической связи. (Множество весов в совокупности обозначается вектором W). Суммирующий блок, соответствующий телу биологического элемента, складывает взвешенные входы алгебраически, создавая выход, который мы будем называть NET.

1.2 Активационные функции

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

$$OUT=1, \text{ если } NET>T$$

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. На рис 2 блок, обозначенный F, принимает сигнал NET и выдает сигнал OUT.

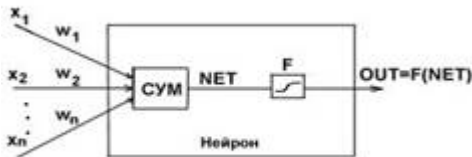


рис.2. Нейрон с активационной функцией

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) функции. Эта функция математически выражается как $F(x)=1/(1+e^{-x})$. Таким образом,

$$OUT=1/(1+e^{-NET})$$

По аналогии с электронными системами активационную функцию можно считать нелинейной усилительной характеристикой искусственного нейрона. Коэффициент усиления вычисляется как отношение приращения величины OUT к вызвавшему его небольшому приращению величины NET.

Он выражается наклоном кривой при определенном уровне возбуждения и изменяется от малых значений при больших отрицательных возбуждениях (кривая почти горизонтальна) до максимального значения при нулевом возбуждении и снова уменьшается, когда возбуждение становится большим положительным. Гроссберг (1973) обнаружил, что подобная нелинейная характеристика решает поставленную им дилемму шумового насыщения. Каким образом одна и та же сеть может обрабатывать как слабые, так и сильные сигналы? Слабые сигналы нуждаются в большом сетевом усилении, чтобы дать пригодный к использованию выходной сигнал. Однако усилительные каскады с большими коэффициентами усиления могут привести к насыщению выхода шумами усилителей (случайными флуктуациями), которые присутствуют в любой физически реализованной сети. Сильные входные сигналы в свою очередь также будут приводить к насыщению

усилительных каскадов, исключая возможность полезного использования выхода. Центральная область логистической функции, имеющая большой коэффициент усиления, решает проблему обработки слабых сигналов, в то время как в области с падающим усилением на положительном и отрицательном концах подходят для больших возбуждений. Таким образом, нейрон функционирует с большим усилением в широком диапазоне уровня входного сигнала.

Рассмотренная простая модель искусственного нейрона игнорирует многие свойства своего биологического двойника. Например, она не принимает во внимание задержки во времени, которые воздействуют на динамику системы. Входные сигналы сразу же порождают выходной сигнал. И что более важно, она не учитывает воздействий функции частотной модуляции или синхронизирующей функции биологического нейрона, которые ряд исследователей считают решающими. Несмотря на эти ограничения, сети, построенные из этих нейронов, обнаруживают свойства, сильно напоминающие биологическую систему. Только время и исследования смогут ответить на вопрос, являются ли подобные совпадения случайными или следствием того, что в модели верно схвачены важнейшие черты биологического нейрона. [1]

1.3 Классификация

- по типу входной информации: Аналоговые нейронные сети (используют информацию в форме действительных чисел); Двоичные нейронные сети (оперируют с информацией, представленной в двоичном виде).
- по характеру обучения: С учителем (выходное пространство решений нейронной сети известно); Без учителя (нейронная сеть формирует выходное пространство решений только на основе входных воздействий). Такие сети называют самоорганизующимися; С критиком (система назначения штрафов и поощрений).
- 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).
- по характеру связей: Сети прямого распространения (Feedforward). Все связи направлены строго от входных нейронов к выходным; Рекуррентные сети. Сигнал с

выходных нейронов или нейронов скрытого слоя частично передается обратно на входы нейронов входного слоя. Как любая система, имеющая обратную связь, рекуррентная сеть стремится к устойчивому состоянию. Как известно, наиболее устойчивое состояние обеспечивается минимизацией энергии системы. Рекуррентная сеть «фильтрует» входные данные, возвращаясь к устойчивому состоянию и, таким образом, позволяет решать задачи компрессии данных и построения ассоциативной памяти; Двухнаправленные сети. В таких сетях между слоями существуют связи как в направлении от входного слоя к выходному, так и в обратном; Самоорганизующиеся карты. Такие сети представляют собой, как правило, двумерную структуру нейронов. Перед обучением структура случайна, нейроны распределены примерно равномерно. При обучении, для каждой обучающей записи рассчитывается точка, которая соответствует ей в структуре сети. Нейрон, находящийся ближе всего к искомой точке, называется нейроном победителем. Веса связей, соединяющих этот нейрон с другими, увеличивается, тем самым несколько упорядочивая структуру. Веса от нейронов, являющихся «соседями» нейрона победителя, к другим нейронам также увеличиваются, но слабее и т. д. Таким образом, чем чаще нейрон «побеждает» при сравнении с признаком, тем «плотнее» к нему находятся другие нейроны. В конце обучения сеть представляет из себя несколько зон сосредоточения нейронов, называемых кластерами. [2]

2. Преимущества нейронных сетей

Круг задач, для решения которых используются нейронные сети, во многом совпадает с задачами, решаемыми традиционными статистическими методами. Поэтому укажем преимущества нейросетей перед несколькими классическими методами статистики.

По сравнению с линейными методами статистики (линейная регрессия, авторегрессия, линейный дискриминант), нейронные сети позволяют эффективно строить нелинейные зависимости, более точно описывающие наборы данных. Из нелинейных методов классической статистики распространен, пожалуй, только байесовский классификатор, строящий квадратичную разделяющую поверхность – нейронная сеть же может построить поверхность более высокого порядка. Высокая нелинейность разделяющей поверхности наивного байесовского классификатора (он не использует ковариационные матрицы классов, как классический байес, а

анализирует локальные плотности вероятности) требует значительного суммарного числа примеров для возможности оценивания вероятностей при каждом сочетании интервалов значений переменных – нейронная сеть же обучается на всей выборке данных, не фрагментируя её, что повышает адекватность настройки нейронной сети.

При построении нелинейных моделей (например, полиномиальных) в статистических программах обычно требуется ручное введение-описание модели в символьном виде с точностью до значений параметров: при $N=10$ независимых переменных полином второй степени будет содержать $N*(N-1)/2=45$ коэффициентов при попарных произведениях переменных, 10 при самих переменных, 10 при квадратах значений переменных, т.е. 65 (66 с учетом неоднородного слагаемого) коэффициентов. При двадцати переменных в выражение войдет уже 231 слагаемое. Вводить такие длинные формулы долго, велик риск опечатки. Нейронная сеть же создается путем указания вида структуры, числа слоев и числа нейронов в каждом слое, что гораздо быстрее. А алгоритмы построения растущих нейросетей и вовсе не требуют первоначального задания размера нейронной сети. Альтернативой нейронной сети при построении сложных нелинейных моделей является только метод группового учета аргументов.

Для сжатия и визуализации данных в статистике разработан метод линейных главных компонент. Нейросети-автоассоциаторы позволяют эффективнее сжимать данные за счет построения нелинейных отображений и визуализировать данные в пространстве меньшего числа нелинейных главных компонент.

По сравнению с методами непараметрической статистики, нейронная сеть с радиальными базисными функциями позволяет сокращать число ядер, оптимизировать координаты и размытость каждого ядра. Это позволяет при сохранении парадигмы локальной ядерной аппроксимации ускорять дальнейший процесс принятия решения.

При обучении нейронной сети вместо критерия качества в виде наименьших квадратов можно использовать робастные критерии, дополнительно вести оптимизацию и других свойств нейронной сети (например, добавляя критерии регуляризации решения или оптимизации структуры нейронной сети). Алгоритмы обучения нейронной сети при этом остаются неизменными.

Необходимость решения прямой и обратной задач обычно требует построения двух моделей. При использовании же нейронных сетей можно обойтись одной сетью, обученной решать прямую задачу.

3. Использование нейронных сетей

Нейронные сети могут решать широкий круг задач обработки и анализа данных – распознавание и классификация образов, прогнозирование, управление и т.д. Конкурентами являются классические методы анализа данных: методы статистики, идентификации систем и управления – частично это обрисовано при обсуждении преимуществ нейронных сетей.

3.1 Требования к компетенции пользователя

Под пользователем здесь понимается человек, непосредственно разрабатывающий нейросетевые модели, а не конечный пользователь готового нейросетевого "решателя". Несмотря на многочисленные заявления о том, что нейронные сети доступны пользователям-неспециалистам, реальное положение вещей на самом деле иное. Во-первых, правильную формулировку задачи никакой компьютер за пользователя не сделает, причем под "правильной формулировкой" понимается не только правильность смысловой постановки задачи, но и правильный выбор математического метода решения и его настроек – а это в случае нейронных сетей подразумевает выбор адекватной структуры нейросети, алгоритма обучения, критерия качества решения задачи и и т.д. Конечно, значительная гибкость и универсальность нейроалгоритмов допускает применение по принципу забивания гвоздей микроскопом, но не всегда это приводит к хорошему результату. Во-вторых, имеющиеся автоматические схемы подбора оптимальных настроек нейрометодов не могут найти правильные решения для более-менее сложных задач. Например, соответствующие процедуры в Statistica Neural Networks не умеют хорошо работать с временными рядами, поскольку не используют приведения ряда к стационарному виду. Так что для пользователей-неспециалистов хороших широкопрофильных автоматических инструментов нет. А если делать вручную, то для получения хороших результатов (я остаюсь пока в рамках примера с временными рядами) придется осваивать как теорию статистического прогнозирования временных рядов и соответствующий модуль пакета Statistica

(или альтернативную статпрограмму), так и нейросетевые методы и программы. И для иных классов задач грамотное применение нейросетей требует хороших знаний как методов нейроинформатики, так и других методов обработки и анализа данных (статистики, например).

3.2 Области применения нейронных сетей

Наверно, в каждой предметной области при ближайшем рассмотрении можно найти постановки нейросетевых задач. Вот список отдельных областей, где решение такого рода задач имеет практическое значение уже сейчас.

Экономика и бизнес: предсказание рынков, автоматический трейдинг, оценка рисков невозврата кредитов, предсказание банкротств, оценка стоимости недвижимости, выявление пере- и недооцененных компаний, автоматическое рейтингование, оптимизация товарных и денежных потоков, автоматическое считывание и распознавание чеков и документов, безопасность транзакций по пластиковым картам.

Медицина: постановка диагноза, обработка медицинских изображений, мониторинг состояния пациента, факторный анализ эффективности лечения, очистка показаний приборов от шумов.

Авионика: обучаемые автопилоты, распознавание сигналов радаров, адаптивное пилотирование сильно поврежденного самолета, беспилотные летательные аппараты.

Связь: сжатие видеоинформации, быстрое кодирование-декодирование, оптимизация сотовых сетей и схем маршрутизации пакетов.

Интернет: ассоциативный поиск информации, электронные секретари и агенты пользователя в Сети, фильтрация информации, блокировка спама, автоматическая рубрикация новостевых лент, адресные реклама и маркетинг для электронной торговли.

Автоматизация производства: оптимизация режимов производственного процесса, контроль качества продукции, мониторинг и визуализация многомерной диспетчерской информации, предупреждение аварийных ситуаций, робототехника.

Политологические и социологические технологии: предсказание результатов выборов, анализ социологических опросов, предсказание динамики рейтингов, выявление значимых факторов, объективная кластеризация электората, исследование и визуализация социальной динамики населения.

Безопасность и охранные системы: идентификация личности по отпечаткам пальцев, голосу, подписи, лицу, распознавание голоса, лиц в толпе, распознавание автомобильных номеров, анализ аэрокосмических снимков, мониторинг информационных потоков в компьютерной сети и обнаружение вторжений, обнаружение подделок.

Ввод и обработка информации: распознавание и обработка рукописных чеков, платежных, иных финансовых и бухгалтерских документов.

Геологоразведка: анализ сейсмических данных, ассоциативные методики поиска полезных ископаемых, оценка ресурсов месторождений.

Обилие приведенных выше применений нейронных сетей - не рекламный трюк. Просто нейросети - это новый, гибкий и мощный инструмент решения разнообразных задач обработки и анализа данных. [3]

Заключение

Развитие нейронных сетей вызвало немало энтузиазма и критики. Некоторые сравнительные исследования оказались оптимистичными, другие - пессимистичными. Для многих задач, таких как распознавание образов, пока не создано доминирующих подходов. Нужно пытаться понять возможности, предпосылки и область применения различных подходов и максимально использовать их дополнительные преимущества для дальнейшего развития интеллектуальных систем.

Множество надежд в отношении нейронных сетей сегодня связывают именно с аппаратными реализациями, но пока время их массового выхода на рынок, видимо, еще не пришло. Они или выпускаются в составе специализированных устройств, или достаточно дороги, а зачастую и то и другое. На их разработку тратится значительное время, за которое программные реализации на самых последних компьютерах оказываются лишь на порядок менее производительными, что делает использование нейропроцессоров нерентабельным. Но все это только вопрос

времени — нейронным сетям предстоит пройти тот же путь, по которому еще совсем недавно развивались компьютеры, увеличивая свои возможности и производительность, захватывая новые сферы применения по мере возникновения новых задач и развития технической основы для их разработки.

Сегодня нейронные сети используются для работы в относительно узких областях, и неизвестно, доверят ли им когда-нибудь решение вопросов, которые требуют понимания социального контекста. Между тем нейронные сети уверенно продолжают проникать в нашу жизнь, и примеров тому немало.

Список использованной литературы

- 1) www.fos.ru./pedagog/9363.html
- 2) www.nekata.ru/index.php?show_section=155
- 3) www.neuropro.ru/neu4.shtml