Содержание

Введение

Основные понятия

Области применения нейронных сетей

Обратная Связь

Классификация нейронных сетей

Структура и принципы работы нейронной сети

Представление знаний

Обучение нейронной сети

Искусственный интеллект и нейронные сети

Введение

Нейронные сети, или, точнее, искусственные нейронные сети, представляют собой технологию, уходящую корнями во множество дисциплин: нейрофизиологию, математику, статистику, физику, компьютерные науки и технику. Они находят свое применение в таких разнородных областях, как моделирование, анализ временных рядов, распознавание образов, обработка сигналов и управление благодаря одному важному свойству — способности обучаться на основе данных при участии учителя или без его вмешательства.

Исследования по искусственным нейронным сетям связаны с тем, что способ обработки информации человеческим мозгом в корне отличается от методов, применяемых обычными цифровыми компьютерами. Мозг представляет собой чрезвычайно сложный, нелинейный, параллельный компьютер. Он обладает способностью организовывать свои структурные компоненты, называемые нейронами , так, чтобы они могли выполнять конкретные задачи (такие как распознавание образов, обработку сигналов органов чувств, моторные функции) во много раз быстрее, чем могут позволить самые быстродействующие современные компьютеры.

Понятие развития нейронов связано с понятием пластичности мозга — способности настройки нервной системы в соответствии с окружающими условиями. Именно пластичность играет самую важную роль в работе нейронов в качестве единиц обработки информации в человеческом мозге. Аналогично, в искусственных нейронных сетях работа проводится с искусственными нейронами. В общем случае нейронная сеть представляет собой машину, моделирующую способ обработки мозгом конкретной задачи. Эта сеть обычно реализуется с помощью электронных компонентов или моделируется программой, выполняемой на цифровом компьютере. Предметом рассмотрения реферата является важный класс нейронных сетей, осуществляющих вычисления с помощью процесса обучения .

Основные понятия

Нейронная сеть — это громадный распределенный параллельный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и предоставляющих их для последующей обработки. Нейронная сеть сходна с мозгом с двух точек зрения.

Знания поступают в нейронную сеть из окружающей среды и используются в процессе обучения.

Для накопления знаний применяются связи между нейронами, называемые синаптическими весами.

Процедура, используемая для процесса обучения, называется алгоритмом обучения. Эта процедура выстраивает в определенном порядке синаптические веса нейронной сети для обеспечения необходимой структуры взаимосвязей нейронов.

Области применения нейронных сетей

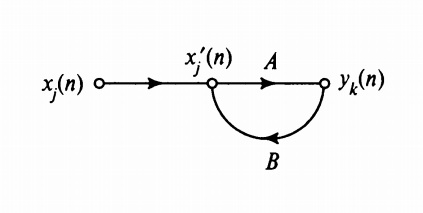
Искусственные нейронные сети в настоящее время широко используются при решении самых разных задач и активно применяются там, где обычные алгоритмические решения оказываются неэффективными или вовсе невозможными. В числе задач, решение которых доверяют искусственным нейронным сетям, можно назвать следующие: распознавание текстов, системы безопасности и видео-наблюдения, автоматизация процессов распознавания образов, адаптивное управление, аппроксимация функционалов, прогнозирование - и это далеко не все. С помощью нейросетей можно выполнять распознавание оптических или звуковых сигналов. Аппаратные реализации ИНС идеально подходят для решения задач идентификации и управления, так как обеспечивают, благодаря параллельной структуре, чрезвычайно высокую скорость выполнения операций.

Описанные возможности в основном относятся к слоистым нейронным сетям, обучаемым алгоритмом обратного распространения, и растущим нейронным сетям на основе вариантов алгоритма каскадной корреляции. Но существуют и другие классы нейронных сетей − нейросети ассоциативной памяти, нейросети для квантования данных, сжатия данных путем построения главных независимых компонент, нейронные сети для разделения смеси сигналов и др. Т.е. круг задач, решаемых нейронными сетями, очень и очень широк, поскольку широк и сам набор нейросетевых алгоритмов.

Обратная Связь

Понятие обратной связи (feedback) характерно для динамических систем, в которых выходной сигнал некоторого элемента системы оказывает влияние на входной сигнала этого элемента. Таким образом, некоторые внешние сигналы усиливаются сигналами, циркулирующими внутри системы.

На самом деле обратная связь присутствует в нервной системе практически любого животного. Более того, она играет важную роль в изучении особого класса нейронных сетей, называемых рекуррентными.



Классификация нейронных сетей

Существует широкий спектр достаточно универсальных способов организации инструментальных средств и собственно процесса применения нейронных сетей на различной программно-аппаратной базе. Всегда можно подобрать наиболее оптимальный для некоторой задачи − всё определяется свойствами задачи и требованиями к решению.

Однако применение нейросетей осложняется рядом причин. Нельзя придумать какую то одну универсальную ИНС, которая бы подошла для различных типов задач. Нейросети используют в двух вариантах:

Строится нейросеть, решающая определенный класс задач,

Под каждый экземпляр задачи строится некоторая нейросеть, находящая квази-оптимальное решение этой задачи.

Существуют несколько видов нейросетей. Их классификация представлена на рисунке 1.1



Рисунок 1.1 Классификация ИНС

Наиболее распространенным семейством сетей прямого действия являются многослойные персептроны, в них нейроны расположены слоями и соединены однонаправленными связями, идущими от входа к выходу сети. Сети прямого действия являются статическими в том смысле, что на заданный вход они вырабатывают одну совокупность выходных значений, не зависящих от предыдущего состояния сети.

Рекуррентные сети являются динамическими, так как в силу обратных связей в них модифицируются входы нейронов, что приводи к изменению состояния сети. Поведение рекуррентных сетей описывается дифференциальными или разностными уравнениями, как правило, первого порядка. Это гораздо расширяет области применения нейросетей и способы их обучения. Сеть организована так, что каждый нейрон получает входную информацию от других нейронов, возможно, и от самого себя, и от окружающей среды.

Так же можно выделить два основных подхода к реализации нейросетей: цифровой и аналоговый. Преимуществом аналоговых реализаций являются: высокое быстродействие, надежность и экономичность. Однако сфера возможного массового использования обучаемых аналоговых нейрочипов достаточно узка. Это обусловлено большой сложностью аппаратной реализации высокоэффективных обучающих алгоритмов и необходимостью специальной подготовки потенциальных пользователей для оптимальной организации адаптивного процесса. В то же время широкое распространение могут получить обученные аналоговые нейрокомпьютеры (нейросети) с фиксированной или незначительно подстраиваемой структурой связей - нейропроцессоры.

Задача создания нейропроцессоров сводится к обучению цифровой нейросетевой модели нужному поведению на обычном цифровом компьютере.

Сети также можно классифицировать по числу слоев. В этом случае важную роль играет нелинейность активационной функции, так как, если бы она не обладала данным свойством или не входила в алгоритм работы каждого нейрона, результат функционирования любой n-слойной нейронной сети сводился бы к перемножению входного вектора сигналов φ на матрицу весовых коэффициентов. То есть фактически такая нейронная сеть эквивалентна однослойной нейросети с весовой матрицей единственного слоя W. Кроме того, нелинейность иногда вводится и в синаптические связи.

Структура и принципы работы нейронной сети

В качестве модели нейрона был выбран бинарный пороговый элемент, вычисляющий взвешенную сумму входных сигналов и формирующий на выходе сигнал величины 1, если эта сумма превышает определенное пороговое значение, и 0 - в противном случае. К настоящему времени данная модель не претерпела серьезных изменений. Были введены новые виды активационных функций. Структурная модель технического нейрона представлена на рисунке 1.3



Рисунок 1.3 Формальная модель искусственного нейрона

На вход искусственного нейрона поступает некоторое множество сигналов, каждый из которых является выходом другого нейрона, или входным сигналом нейросетевой модели. Каждый вход умножается на соответствующий вес, аналогичный синаптической силе биологического нейрона. Вес определяет, насколько соответствующий вход нейрона влияет на его состояние. Все произведения суммируются, определяя уровень активации нейрона s. Состояние нейрона определяется по формуле.

, (1.1)

где φ - множество сигналов, поступающих на вход нейрона,

wi - весовые коэффициенты нейрона.

Далее сигнал s преобразуется активационной (передаточной) функцией нейрона F в выходной сигнал y. Математически это можно выразить формулой:

, (1.2)

где n - размерность вектора входов,

w0 - "нейронное смещение", вводимое для инициализации сети, - подключается к неизменяемому входу +1,F - активационная функция нейрона.

Нейроны могут группироваться в сетевую структуру различным образом.

Функциональные особенности нейронов и способ их объединения в сетевую структуру определяет особенности нейросети. Для решения задач идентификации и управления наиболее адекватными являются многослойные нейронные сети (МНС) прямого действия или многослойные персептроны. При проектировании МНС нейроны объединяют в слои, каждый из которых обрабатывает вектор сигналов от предыдущего слоя.

Минимальной реализацией является двухслойная нейронная сеть, состоящая из входного (распределительного), промежуточного (скрытого) и выходного слоя.



Рисунок 1.4 Структурная схема двухслойной нейронной сети.

Реализация модели двухслойной нейронной сети прямого действия имеет следующее математическое представление:

, (1.7)

где nφ - размерность вектора входов φ нейронной сети;

nh - число нейронов в скрытом слое;

θ - вектор настраиваемых параметров нейронной сети, включающий весовые коэффициениы и нейронные смещения (wji, Wij)

fj (x) - активационная функция нейронов скрытого слоя;

Fi (x) - активационная функция нейронов выходного слоя.

Персептрон представляет собой сеть, состоящую из нескольких последовательно соединенных слоев формальных нейронов (рисунок 1.3). На низшем уровне иерархии находится входной слой, состоящий из сенсорных элементов, задачей которого является только прием и распространение по сети входной информации. Далее имеются один или, реже, несколько скрытых слоев. Каждый нейрон на скрытом слое имеет несколько входов, соединенных с выходами нейронов предыдущего слоя или непосредственно со входными сенсорами φ1. φn, и один выход. Нейрон характеризуется уникальным вектором настраиваемых параметров θ. Функция нейрона состоит в вычислении взвешенной суммы его входов с дальнейшим нелинейным преобразованием ее в выходной сигнал.

Представление знаний

Под знанием понимается хранимая информация или модели, используемые человеком или машиной для интерпретации, предсказания и реакции на внешние события. Знания о мире включают два типа информации.

Известное состояние окружающего мира, представленное имеющимися в наличии достоверными фактами. Такая информация называется априорной.

Наблюдения за окружающим миром, полученные с помощью сенсоров, адаптированных для конкретных условий, в которых должна функционировать данная нейронная сеть.

Примеры могут быть маркированными и немаркированными. В маркированных примерах входному сигналу соответствует желаемый отклик. Немаркированные примеры состоят из нескольких различных реализаций одного входного сигнала. В любом случае набор примеров, будь то маркированных или нет, представляет собой знания об интересующей предметной области, на основании которых и проводится обучение нейронной сети.

Вопрос представления знаний в нейронной сети является очень сложным. Тем не менее можно выделить четыре общих правила.

Сходные входные сигналы от схожих классов должны формировать единое представление в нейронной сети. Исходя из этого, они должны быть классифицированы как принадлежащие к одной категории.

Элементы, отнесенные к различным классам, должны иметь в сети как можно более отличные представления. Это правило прямо противоположно первому.

Если некоторое свойство имеет важное значение, то для его представления в сети необходимо использовать большое количество нейронов.

В структуру нейронной сети должны быть встроены априорная информация и инварианты, что упрощает архитектуру сети и процесс ее обучения.

Это правило играет особую роль, поскольку правильная конфигурация сети обеспечивает ее специализацию. Нейронная сеть со специализированной структурой обычно включает значительно меньшее количество свободных параметров. При этом на обучение затрачивается меньше времени, и такая сеть обладает лучшей обобщающей и пропускной способностями и их стоимость создания сокращается. Чтобы встроить априорную информацию в структуру нейронной сети можно использовать комбинацию двух следующих приемов:

Ограничение сетевой архитектуры с помощью локальных связей, называемых рецепторными полями,

Ограничение выбора синаптических весов за счет совместного использования весов.

Для того чтобы создать систему распознавания объекта, речи или эхолокации, учитывающую явления такого рода, необходимо принимать во внимание диапазон трансформаций наблюдаемого сигнала. Соответственно основным требованием при распознавании образов является создание такого классификатора, который инвариантен к этим трансформациям.

Существуют как минимум три приема обеспечения инвариантности нейронной сети классификации к подобным трансформациям.

Структурная инвариантность: Синаптические связи между отдельными нейронами сети строятся таким образом, чтобы трансформированные версии одного и того же сигнала вызывали один и тот же выходной сигнал.

Инвариантность по обучению: Сеть обучается на множестве примеров одного и того же объекта, при этом в каждом примере объект подается в несколько измененном виде (например, снимки с разных ракурсов). Если количество таких примеров достаточно велико и если нейронная сеть обучена отличать разные точки зрения на объект, можно ожидать, что эти данные будут обобщены и сеть сможет распознавать ракурсы объекта, которые не использовались при обучении.

Использование инвариантных признаков: Этот метод основывается на предположении, что из входного сигнала можно выделить информативные признаки, которые описывают самую существенную информацию, содержащуюся в наборе данных, и при этом инвариантны к трансформациям входного сигнала.



Обучение нейронной сети

Следующий этап создания нейросети - это обучение. Способность к обучению является основным свойством мозга. Для искусственных нейронных сетей под обучением понимается процесс настройки архитектуры сети (структуры связей между нейронами) и весов синаптических связей (влияющих на сигналы коэффициентов) для эффективного решения поставленной задачи. Обычно обучение нейронной сети осуществляется на некоторой выборке. По мере процесса обучения, который происходит по некоторому алгоритму, сеть должна все лучше и лучше (правильнее) реагировать на входные сигналы.

Выделяют три типа обучения: с учителем, самообучение и смешанный. В первом способе известны правильные ответы к каждому входному примеру, а веса подстраиваются так, чтобы минимизировать ошибку. Обучение без учителя позволяет распределить образцы по категориям за счет раскрытия внутренней структуры и природы данных, выходы НС формируются самостоятельно, а веса изменяются по алгоритму, учитывающему только входные и производные от них сигналы. При смешанном обучении комбинируются два вышеизложенных подхода.

Поскольку ошибка зависит от весов нелинейно, получить решение в аналитической форме невозможно, и поиск глобального минимума осуществляется посредством итерационного процесса - так называемого обучающего алгоритма. Разработано уже более сотни разных обучающих алгоритмов, отличающихся друг от друга стратегией оптимизации и критерием ошибок. Обычно в качестве меры погрешности берется средняя квадратичная ошибка (СКО):

 (1.8)

где М - число примеров в обучающем множестве;

d - требуемый выходной сигнал;

y - полученный выходной сигнал.

Обучение нейросети производится методом градиентного спуска, т.е. на каждой итерации изменение веса производится по формуле.

, (1.9)

где ε - коэффициент обучения, определяющий скорость обучения.

Отметим два свойства полной ошибки. Во-первых, ошибка E=E (W) является функцией состояния W, определенной на пространстве состояний. По определению, она принимает неотрицательные значения. Во-вторых, в некотором обученном состоянии W\*, в котором сеть не делает ошибок на обучающей выборке, данная функция принимает нулевое значение. Следовательно, обученные состояния являются точками минимума введенной функции E (W).

Таким образом, задача обучения нейронной сети является задачей поиска минимума функции ошибки в пространстве состояний

Искусственный интеллект и нейронные сети

Основной задачей искусственного интеллекта является разработка парадигм или алгоритмов, обеспечивающих компьютерное решение когнитивных задач, свойственных человеческому мозгу.

Системы искусственного интеллекта должны обеспечивать решение следующих трех задач: накопление знаний, применение накопленных знаний для решения проблемы и извлечение знаний из опыта. Системы искусственного интеллекта реализуют три ключевые функции:

Представление: Одной из отличительных черт систем искусственного интеллекта является использование символьного языка для представления общих знаний о предметной области и конкретных знаний о способах решения задачи.

Рассуждения : Под рассуждениями обычно понимается способность решать задачи. Для того чтобы систему можно было назвать разумной, она должна описывать и решать широкий спектр задач, понимать явную и неявную информацию, иметь механизм управления , определяющий операции, выполняемые для решения отдельных задач.

Обучение: В простейшей модели машинного обучения информацию для обучаемого элемента предоставляет сама среда. Обучаемый элемент использует полученную информацию для модернизации базы знаний, знания из которой функциональный элемент затем использует для выполнения поставленной задачи.

