

Матвеева С.С.

«Искусственные нейронные сети
для школьников»

часть I

«Знакомство с ИИ
и нейросетями»

учебное пособие
для внеурочных занятий

Аннотация

Учебное пособие «Искусственные нейронные сети для школьников» является отдельным курсом для изучения IT направления в области создания нейронных сетей и рассчитано на учащихся средней и старшей школы. Изучение курса позволяет получить знания о искусственном интеллекте и нейронных сетях, сфере их применения, освоить первичные навыки по программированию и обучению простейшей нейросети.

Учебное пособие «Искусственные нейронные сети для школьников» рассчитан на среднюю и старшую школу и делится на три части:

- **1 часть** «Знакомство с ИИ и нейросетями»;
- **2 часть** «Обучение нейронных сетей, профессия нейропрограммист»;
- **3 часть** «Обучение и тестирование нейросети, профессия тестировщик».

Каждое пособие может быть самостоятельным курсом и использоваться для конкретных возрастных групп. Знания и умения полученные при изучении курса, помогут определиться с профессией будущего.

Данное учебное пособие является 1 частью. **1 Часть «Знакомство с ИИ и нейросетями»** это быстрый старт в теорию ИНС. Знакомство с сетями, их обучением и работой.

Электронный вариант учебного пособия proneyroset.ru

Введение

Уважаемые школьники!

Мы живем во время стремительного развития технологий, уже сейчас на службе у человека есть различные роботы, выполняющие сложную и трудоемкую работу. Роботами управляет человек с помощью программного обеспечения и операторов, отслеживающих четкое выполнение поставленных задач.

Современные технологии развиваются настолько быстро, что в скором будущем роботы сами без контроля человека, смогут решать задачи, обучаться, развиваться, «думать», приносить пользу в различных сферах жизни: образовании, медицине, промышленности, в творчестве и т.д. Для этого роботам будет нужен искусственный «мозг». Искусственный мозг имеет название **Искусственная Нейронная Сеть (ИНС)**-это одно из направлений в разработке **ИИ (Искусственный интеллект)**.

ИНС – является упрощенной моделью биологической нейронной сети, представляющей собой совокупность искусственных нейронов, взаимодействующих между собой. Созданием ИНС занимаются нейропрограммисты. От нейропрограммистов зависит какими роботы будущего будут, что они будут уметь делать, будут ли они безопасны для человека, будут ли они жить с нами бок о бок в повседневной жизни.

При изучении курса «Искусственные нейронные сети для школьников» Вы узнаете направления и перспективы развития ИИ, технологии развивающие его, сферы использования, результаты применения, познакомитесь с профессией нейропрограммист.

ВСЕМ ЖЕЛАЮ УСПЕХА В ИЗУЧЕНИИ РАБОТЫ НЕЙРОННЫХ СЕТЕЙ!

I Часть

Знакомство с ИИ и нейросетями.

§ 1.

Знакомство с искусственным интеллектом (ИИ)

Словарь

- Искусственный интеллект
- Нейросети
- Машинное обучение
- Слабый ИИ
- Сильный ИИ
- Большие данные (Big Data)

Искусственный интеллект – это свойство интеллектуальной системы выполнять те функции и задачи, которые обычно характерны для разумных существ. Это может быть проявление каких-то творческих способностей, склонность к рассуждению, обобщение, обучение на основании полученного ранее опыта и так далее.

Термин artificial intelligence (с английского переводится как «искусственный интеллект») был упомянут в 1956 году Джоном МакКарти, основателем функционального программирования и изобретателем языка Lisp, на конференции в Университете Дартмута.

Алан Тьюринг¹ в 1950 году, предложил считать интеллектуальными те системы, которые в общении не будут отличаться от человека.

Тест Тьюринга-человек обменивается текстовыми сообщениями с двумя собеседниками, находящимися в комнате. Переписка проводится через контролируемые интервалы времени. Один собеседник –человек, второй-компьютер. Если судья не сможет определенно сказать, кто из собеседников является человеком, то тест пройден. На данный момент тест не прошла ни одна из машин.

Тест Тьюринга. Рис.1



Самая ранняя успешная программа искусственного интеллекта была создана Кристофером Стрейчи в 1951 году. А уже в 1952 году она играла в шашки с человеком и удивляла зрителей своими способностями предсказывать ходы. По этому поводу в 1953 году Тьюринг опубликовал статью о шахматном программировании.

В 1965 году специалист Массачусетского технологического университета Джозеф Вайценбаум разработал программу «Элиза», которая ныне считается прообразом современной Siri. В 1973 году была изобретена «Стэндфордская тележка», первый беспилотный автомобиль, контролируемый компьютером. К концу 1970-х интерес к ИИ начал спадать.

Новое развитие искусственный интеллект получил в середине 1990-х. Самый известный пример – суперкомпьютер IBM Deep Blue, который в 1997 году обыграл в шахматы чемпиона мира Гарри Каспарова. Сегодня подобные сети развиваются очень быстро за счет цифровизации информации, увеличения ее оборота и объема. Машины довольно быстро анализируют информацию и обучаются, впоследствии они действительно приобретают способности, ранее считавшиеся чисто человеческой прерогативой.

Отличие ИИ от нейросетей и машинного обучения

Нейросети представляют собой математическую модель, компьютерный алгоритм, работа которого основана на множестве искусственных нейронов. Суть этой системы в том, что ее не нужно заранее программировать. Она моделирует работу нейронов человеческого мозга, проводит элементарные вычисления и обучается на основании предыдущего опыта, но это не соотносимо с ИИ.

Искусственный интеллект, как мы помним, является свойством сложных систем выполнять задачи, обычно свойственные человеку. К ИИ часто относят узкоспециализированные компьютерные программы, также различные научно-технологические методы и решения. ИИ в своей работе имитирует человеческий мозг, при этом основывается на прочих логических и математических алгоритмах или инструментах, в том числе нейронных сетях.

Под **машинным обучением** понимают использование различных технологий для самообучающихся программ. Соответственно, это одно из многочисленных направлений ИИ. Системы, основанные на машинном обучении, получают базовые данные, анализируют их, затем на основе полученных выводов находят закономерности в сложных задачах со множеством параметров и дают точные ответы. **Один из наиболее распространенных вариантов организации машинного обучения – применение нейросетей.**

Вывод: Если сравнивать с человеком, то ИИ подобен головному мозгу, машинное обучение – это один из многочисленных способов обработки поступающих данных и решения назревающих задач, а нейросети соответствуют объединению более мелких, базовых элементов мозга – нейронов.

Разница между искусственным и естественным интеллектом

Сравнивать искусственный и естественный интеллект можно лишь по некоторым общим параметрам. Например, человеческий мозг и **компьютер работают по примерно схожему принципу, включающему четыре этапа – кодирование, хранение данных, анализ и предоставление результатов.** И естественный, и искусственный разум склонны к самообучению, они решают те или иные задачи и проблемы, используя специальные алгоритмы.

Помимо общих умственных способностей к рассуждению, обучению и решению проблем, человеческое мышление также имеет эмоциональную окраску и сильно зависит от влияния социума. Искусственный интеллект не имеет никакого эмоционального характера и не ориентирован социально.

Если говорить об IQ – большинство ученых склонны считать, что сей параметр оценки никак не связан с искусственным интеллектом. С одной стороны, это действительно так, ведь стандартные IQ-тесты направлены на измерение «качества» человеческого мышления и связаны с развитием интеллекта на разных возрастных этапах.

С другой стороны, для ИИ создан собственный «IQ-тест», названный в честь Тьюринга. Он помогает определить, насколько хорошо машина обучилась и способна ли она уподобиться в общении человеку. Это своего рода планка для ИИ, установленная людьми. А ведь все больше ученых склоняется к тому, что скоро компьютеры обгонят человечество по всем параметрам. Развитие технологий идет по непредсказуемому сценарию, и вполне допустимо, что так и будет.

Применение ИИ в современной жизни

В зависимости от области и обширности сферы применения, выделяют два вида ИИ:

Слабый ИИ - Weak AI, называемый еще «слабым», решает узкоспециализированные задачи – диагностика в медицине, управление роботами, работа на базе электронных торговых платформ;

Сильный ИИ- Strong AI, «сильный» решение глобальных задач.

Так, одна из наиболее популярных сфер применения ИИ – это Big Data (большие данные) в коммерции. Крупные торговые площадки используют подобные технологии для исследования потребительского поведения. Компания «Яндекс» вообще создает с их помощью музыку. В некоторые мобильные приложения встроены голосовые помощники вроде Алисы. Они упрощают процесс навигации и совершения покупок в сервисе. И не стоит забывать про программы с нейросетями, обрабатывающими фото и видео.

ИИ также внедряют в производственные процессы для фиксации действий работников. Не обошлось и без внедрения новых технологических решений в транспортной сфере. Так, искусственный интеллект мониторит состояние на дорогах, фиксирует пробки, обнаруживает разные объекты в неположенных местах. А про

автономное (беспилотное) вождение и так постоянно говорят.

Популярные бренды внедряют ИИ в свои системы для анализа потребностей клиентов. Стремительно развивается использование подобных систем в системах здравоохранения, в основном при диагностике заболеваний, разработке лекарств, создании медицинских страховок, проведении клинических исследований и так далее.

Банки и государственные сервисы используют чат-ботов, голосовых помощников, например на сервисе «Госуслуг» с 31 марта 2021 года запущен робот помощник Макс.

Перечислить разом все области, в которых задействован искусственный интеллект, практически нереально. На данный момент он затрагивает все больше самых разных сфер. И причин на то немало – та же автоматизация производственных процессов, стремительный рост информационного оборота и инвестиций в эту сферу, даже социальное давление.

Влияние на различные области

ИИ все больше проникает в экономическую сферу, и, по некоторым прогнозам, это позволит увеличить объем глобального рынка на 15,7 трлн долларов к 2030 году. Лидирующую позицию в освоении сей технологии занимают США и Китай, однако некоторые развитые страны вроде Канады, Сингапура, Германии и Японии не отстают. В России в июле 2018 г. заработала Единая биометрическая система (ЕБС).

Некоторые ученые отмечают риски внедрения ИИ в повседневную жизнь. Так, британский ученый Стивен Хокинг² считал, что создать ИИ, превосходящий человека по всем параметрам, все же удастся, но справиться с ним будет нам не под силу, и людям будет нанесен существенный вред. Илон Маск³ же считает, что искусственный разум в дальнейшем будет нести куда большую угрозу по сравнению с ядерным оружием.

Перспективы развития искусственного интеллекта

Современные компьютеры приобретают все больше знаний и «умений». Скептики же утверждают, что все возможности ИИ – не более чем компьютерная программа, а не пример самообучения. Однако это не мешает технологии широко распространяться в самых различных сферах и открывать невиданные ранее потенциалы для развития. Со временем компьютеры будут становиться все мощнее, а ИИ еще быстрее совершенствоваться в своем развитии.

Алан Мэтисон Тьюринг¹, ОВЕ — английский математик, логик, криптограф, оказавший существенное влияние на развитие информатики.

Стивен Уильям Хокинг² (англ. Stephen William Hawking; 8 января 1942, Оксфорд, Великобритания — 14 марта 2018, Кембридж, Великобритания) — английский физик-теоретик, космолог и астрофизик, писатель.

Илон Рив Маск³ — американский предприниматель, инженер и миллиардер. Основатель, генеральный директор и главный инженер компании SpaceX

§ 2.

Биологический нейрон и биологическая нейронная сеть

Словарь

- Нейрон
- Аксон
- Дендрит
- Синописис
- Структура нейрона
- Биологическая нейронная сеть

Для того чтобы создавать и обучать искусственные нейронные сети, необходимо понимать что такое биологический нейрон, знать его строение, структуру. Иметь представление о биологической нейронной сети.

Нейрон или нервная клетка-это электрически возбудимая клетка, которая общается с другими клетками через специализированные соединения, называемые синапсами.

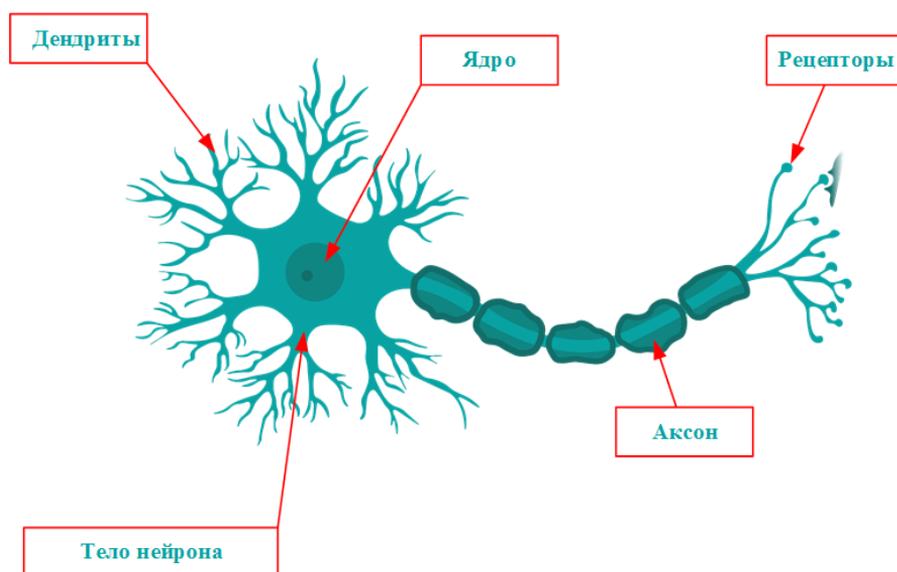
Части нейрона:

Тело нейрона - представляет собой скопление цитоплазмы, в которой располагается крупное круглое ядро (в нервных клетках вегетативной нервной системы встречается по 2–3 ядра). Ядро регулирует синтез белков и несет в себе генетическую информацию

Аксон – отросток по которому импульс идет от тела нейрона на периферию (к другому нейрону или к исполнительной клетке)

Дендрит-отросток по которому импульс идет к тела нейрона с периферии (от другого нейрона или рецептора)

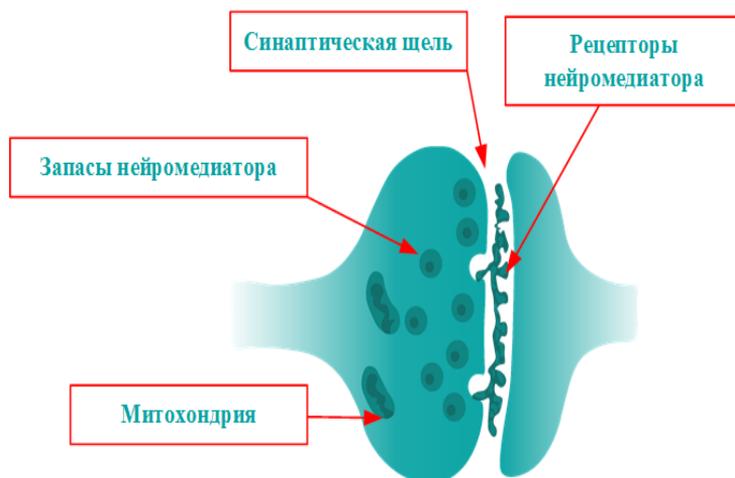
Биологический нейрон. Рис. 1.



Нейрон способен пропускать нервный импульс только в одном направлении – от дендрита через тело клетки к аксону. Дендриты принимают импульс от Аксона другого нейрона в месте передачи импульса образуется **Синапс**.

Синапс - место контакта между двумя нейронами или между нейроном и получающей сигнал эффекторной клеткой. Служит для передачи нервного импульса между двумя клетками.

Устройство синапсиса. Рис. 2.



Чем важнее сигнал-тем больше нейромедиатора.

Биологическая нейронная сеть — сеть, состоящая из биологических нейронов, которые связаны или функционально объединены в нервной системе.

Вывод: Дендриты-принимают сигналы. Аксон-отдает сигнал. Синапс-место соединения двух нейронов. Информация передается между нейронами с помощью синапсов. Чем важнее сигнал, тем больше нейромедиатора передается от нейрона к нейрону.

Вопросы:

1. Как проходит нервный импульс в нейроне?
2. Что представляет из себя структура нейрона?
3. Какую роль играют Дендриты и Аксоны?
4. Какую роль выполняет Синапс?
5. В каком случае от нейрона к нейрону передается большее количество нейромедиатора?

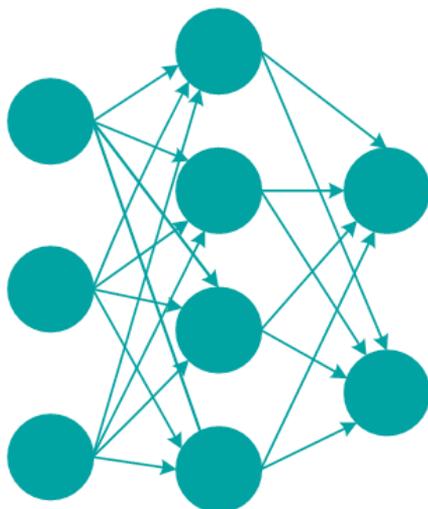
§ 3

Искусственный нейрон

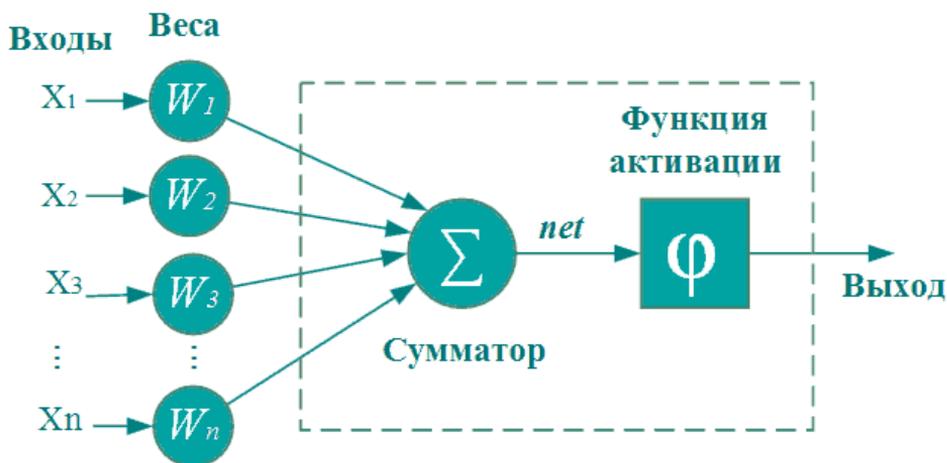
Словарь

- Искусственный нейрон
- Входные сигналы
- Синаптические веса
- Блок суммирования
- Постсинаптическое возбуждение
- Блок нелинейного преобразования
- Выходной сигнал
- Взвешанная сумма

Структура искусственной нейронной сети. Рис. 3.



Модель искусственного нейрона. Рис. 4.



Математическая модель искусственного нейрона

с n входами:

$$\text{out} = \phi \sum_{i=1}^n x_i w_i$$

ϕ – функция активации

n

$\sum_{i=1}^n x_i w_i$ – взвешенная сумма, как сумма n произведений входных сигналов на соответствующие веса

Искусственный нейрон, из которых состоит ИНС, имеет намного более простую структуру: у него есть несколько входов, на которых он принимает различные сигналы, преобразует их и передает другим нейронам.

1. **Входные сигналы** поступают от других искусственных нейронов в интервале от 0 до 1.
2. **Синаптические веса** выступают в роли нейромедиаторов (смотри § 2, определение нейромедиаторов) и умножаются на входящие сигналы.
3. **Взвешенные сигналы** направляются в сумматор, который агрегирует сигналы в взвешенную сумму.
4. **Сигнал на выходе** из сумматора называется **постсинаптическим возбуждением**.
5. Иногда для получения сигнала используют **функцию активации** (блок нелинейного преобразования) **нейрона**. $R^n \rightarrow R$, которая преобразует несколько входных параметров в один выходной.

Вывод: Сам искусственный нейрон ничего делать не умеет, он только принимает поступившие на вход сигналы, которые умножаются на веса (чем больше вес, тем важнее сигнал, тем больше нейромедиатора). Взвешенные сигналы суммируются, преобразовываются и поступают на выход. Поэтому, нам нужен не один нейрон, а сеть нейронов.

Работа сумматора.

Поступившие на входы сигналы умножаются на свои веса. Сигнал первого входа x_1 умножается на соответствующий этому входу вес w_1 . В итоге получаем $x_1 w_1$. И так до n -ого входа. В итоге на последнем входе получаем $x_n w_n$. Теперь все произведения передаются в сумматор. Уже исходя из его названия можно понять, что он делает.

Он просто суммирует все входные сигналы, умноженные на соответствующие веса:

$$x_1 w_1 + x_2 w_2 + \dots + x_n w_n = \sum_{i=1}^n x_i w_i$$

Результатом работы сумматора является число, называемое взвешенной суммой.

Взвешенная сумма (Weighted sum) (**net**) — сумма входных сигналов, умноженных на соответствующие им веса.

$$\text{net} = \sum_{i=1}^n x_i w_i$$

Пример «Покупка пончика».

У нас есть один искусственный нейрон, его задача определить стоит ли покупать пончик.

На входы нейрона подаем входные данные:

1. Низкая цена пончика;
2. Хорошее качество муки;
3. Много глазировки;
4. Пончик выпекали утром.

Входные параметры необходимо охарактеризовать от **0** до **1**. Если например цена низкая на пончик подаем на этот вход параметр равный **1**. Со всеми остальными параметрами поступаем аналогично.

У нашего нейрона **4** входных параметра, то есть **4** входа, поэтому мы выставляем **4** весовых коэффициента. Вес можно определить в зависимости от важности каждого параметра. Чем выше коэффициент, тем больше вес параметра.

1. 1
2. **4**
3. 1
4. **5**

Мы видим, что важно приобрести свежий пончик из качественной муки, а количество глазировки и цена в нашем примере роль не играет.

Четвертый и второй входы:

«Пончик выпекали утром» и «Хорошее качество муки» будут решающие при принятии нейроном решения. Значит, четвертый и второй входы будут играть решающую роль при выборе покупки пончика.

Пусть на входы нейрона мы подаем следующие сигналы:

1. 1
2. 0
3. 0
4. 1

Умножаем веса входов на сигналы соответствующих входов:

1. 1 (1*1)
2. 0 (0***4**)
3. 0 (0*1)
4. 5 (1***5**)

Взвешенная сумма для такого набора входных сигналов равна 6:

$$\text{net} = \sum_{i=1}^n x_i w_i = 1 + 0 + 0 + 5 = 6$$

Получив взвешанную сумму, нейрон может вывести ее на выход, но она будет бессмысленна, нейрону нужно обработать результат и сформировать адекватный выходной сигнал.

Для того чтобы нейрон смог обработать сигнал и подать его на выход, то есть выбрать покупать пончик или нет, необходимо использовать **функцию Активации**.

Вопросы:

1. Что такое искусственный нейрон?
2. В каком интервале поступают входные сигналы от других искусственных нейронов?
3. Что делает сумматор?
4. Что такое постсинаптическое возбуждение?
5. Что преобразует функция активации?
6. Как называется число, которое является результатом работы сумматора?
7. Что такое взвешанная сумма?

Рассмотрите пример «Покупка пончика» ответьте на вопросы.

1. Сколько параметров поступили на вход искусственного нейрона?
2. Какие два входных параметра в примере, наиболее важные при покупке пончика?
3. Какую роль будут играть наиболее важные параметры при принятии решения нейроном?
4. На что умножаем веса входов?
5. Как получили взвешанную сумму равную 6 для имеющегося набора входных сигналов, объясните формулу?

§ 4.

Функция активации. Функция единого скачка.

Словарь

- Функция активации
- Функция единого скачка
- Порог
- График функции единого скачка

Функция активации (Activation function) ($\phi(\mathbf{net})$) — функция, принимающая взвешенную сумму как аргумент. Значение этой функции и является выходом нейрона (**out**).

$$\mathbf{out} = \phi(\mathbf{net})$$

Известные функции активации:

- Функция единого скачка
- Сигмоидальные функции

В этом параграфе мы рассмотрим функцию единого скачка.

Функция единого скачка.

Выхода нейрон всегда может быть равен только либо **0** либо **1**.

Поэтому необходимо использовать порог, например: сумма больше порога, то сигнал на выход равен **1**, если меньше, то **0**.

Допустим порог в нашей задаче (§ 3, пример «Покупка пончика») равен 5., взвешанная сумма равна 6. Следовательно наш порог больше 5, а значит выходной сигнал =1. Следовательно нейрон выбирает покупку пончика. Если бы порог был меньше 5, то пончик куплен бы не был.

Порог b

$$b=6$$

Вывод: Нейрон смотрит на взвешенную сумму и если она получается \geq его порога, то нейрон выдает выходной сигнал равный 1, если $<$ порога выдает сигнал равный 0.

Задание 1 (продолжаем работать с задачей §3 «Покупка пончика»).

Определите будет ли нейроном выдан сигнал о покупке пончика, если на вход были даны следующие сигналы.

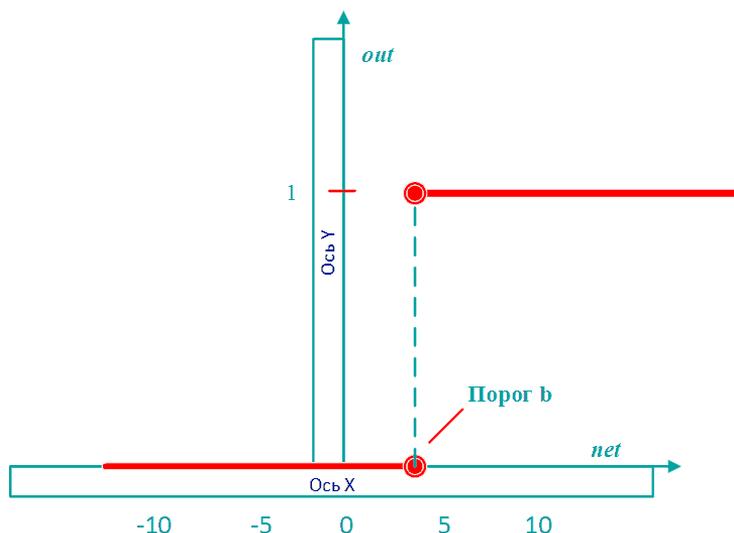
1. 0
2. 1
3. 1
4. 0

Ход решения:

- 1) Найдем взвешенную сумму;
- 2) Найдем чему будет равен сигнал данный нейроном на выход.

График функции единичного скачка. Рис. 5.

Функция единичного скачка



На горизонтальной оси расположены величины взвешенной суммы. На вертикальной оси — значения выходного сигнала. Как легко видеть, возможны только два значения выходного сигнала: 0 или 1. Причем 0 будет выдаваться всегда от минус бесконечности и вплоть до некоторого значения взвешенной суммы, называемого порогом. Если взвешенная сумма равна порогу или больше него, то функция выдает 1. Все предельно просто.

Математическая запись функции активации

(составная функция единичного скачка).

Составная функция – объединение нескольких правил под одной функцией.

- выход нейрона (out)
- взвешенная сумма (net)
- порог (b)

$$\text{out}(\text{net}) = \begin{cases} 0, & \text{net} < b \\ 1, & \text{net} \geq b \end{cases}$$

Вывод: Выход нейрона (out) зависит от взвешенной суммы (net) следующим образом: если net (взвешенная сумма) меньше какого-то порога (b), то out (выход нейрона) равен 0. А если net больше или равен порогу b, то out равен 1.

Вопросы:

1. Опишите функцию единичного скачка.
2. Как выглядит составная функция единичного скачка (формула)?

§ 5.

Функция активации. Сигмоидальные функции.

Словарь

- Логистическая функция
- Гиперболический тангенс

В этом параграфе мы рассмотрим сигмоидальные функции.

Сигмоидальные функции:

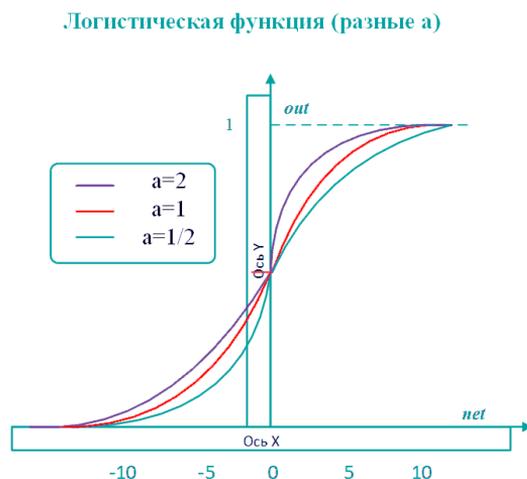
- Логистическая функция
- Гиперболический тангенс

Логистическая функция.

Некоторые из сигмоидальных функций, например **Логистическая** используются в качестве функции активации в искусственных нейронах. Если при использовании функции единого скачка (смотрите §4 функция единого скачка), нейрон на выход подает сигнал равный 1 (покупаем пончик) или 0 (не покупаем пончик), то **при использовании логистической функции можно результат приблизить к точности.** Мы можем использовать приближенные веса в сторону 0 и 1, например 0,8; 0,5; 0,1 и так далее. Причем, чем ближе к 0, тем яснее, что пончик не купим, а чем ближе к 1, яснее что пончик купим. Другими словами, если ряд параметров, которые отвечают за положительный результат имеют большие веса, то и сигнал на выходе, будет максимально приближен к 1 и наоборот.

Например выход нейрона равен 0,7 это значит, что он считает, что купить пончик (смотри задача §3), все-таки стоит. Если выход нейрона равен 0,4 то скорее всего, пончик не будет куплен.

График логистической функции. Рис. 6.



Логическая запись функции активации через логистическую функцию.

$$\phi(\text{net})=1/1+\exp(-a*\text{net})$$

где параметр **a** характеризует степень крутизны функции.

Вспомним наш искусственный нейрон, определяющий, надо ли покупать пончик. В случае с функцией единичного скачка все понятно. Мы либо покупаем пончик (1), либо нет (0).

Здесь же мы до конца не уверены сможем мы купить пончик или нет, но зато при использовании логистической функции в качестве функции активации мы будем получать цифру между 0 и 1. А значит, более взвешанная сумма, будет стремиться на выходе показать 1, менее взвешанная сумма 0. Например выход нашего нейрона равен 0.7. Это говорит о том, что нейрон считает, что купить пончик стоит.

Вывод: логистическая функция будет всегда в пределах от 0 до 1. При ее использовании результатом может быть не только 0 и 1, но и любое число между ними. А это значит, что нейрон может сделать более реальный выбор. Поэтому логистическую функцию в качестве функции активации используют чаще.

Гиперболический тангенс

Гиперболический тангенс - применяется в качестве функции активации биологами для более реалистичной модели нервной клетки.

Такая функция позволяет получить на выходе значения разных знаков (например, от -1 до 1), что может быть полезным для ряда сетей.

Математическая запись функции активации через гиперболический тангенс.

$$\phi(\text{net})= \tanh(\text{net}/a)$$

где параметр **a** также определяет степень крутизны графика функции

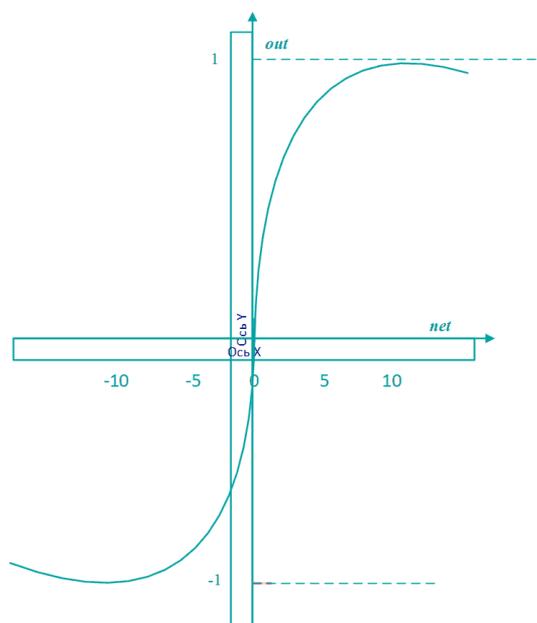
Такая функция позволяет получить на выходе значения разных знаков (например, от -1 до 1), что может быть полезным для ряда сетей.

Гиперболический тангенс обладает всеми полезными свойствами, которые имеет и логистическая функция.

Вывод: У нейрона есть входы. На них подаются сигналы в виде чисел. Каждый вход имеет свой вес (тоже число). Сигналы на входе умножаются на соответствующие веса. Получаем набор «взвешенных» входных сигналов. Далее этот набор попадает в сумматор, который просто складывает все входные сигналы, помноженные на веса. Получившееся число называют взвешенной суммой. Затем взвешенная сумма преобразуется функцией активации и мы получаем выход нейрона.

График гиперболической функции. Рис. 7.

Гиперболический тангенс



Вопросы:

1. Какие виды сигмоидальных функций бывают?
2. Чем отличаются Логистическая функция и гиперболический тангенс?
3. Опишите полный цикл работы искусственного нейрона.

§ 6.

Виды ИНС (Искусственных нейронных сетей)

Словарь

- Однослойные ИНС
- Многослойные ИНС
- Сети прямого распределения

Искусственная нейронная сеть состоит из нейронов, взаимодействующих между собой, однако представляет собой упрощенную модель.

В ИНС есть три основных слоя:

Входной слой-искусственных нейронов принимает сигналы и распределяет их остальным нейронам. Нейроны этого слоя не производят никаких вычислений.

Скрытый слой-обрабатывает, полученную в виде сигналов, информацию.

Выходной слой-получаем на выходе готовый результат работы нейросети.

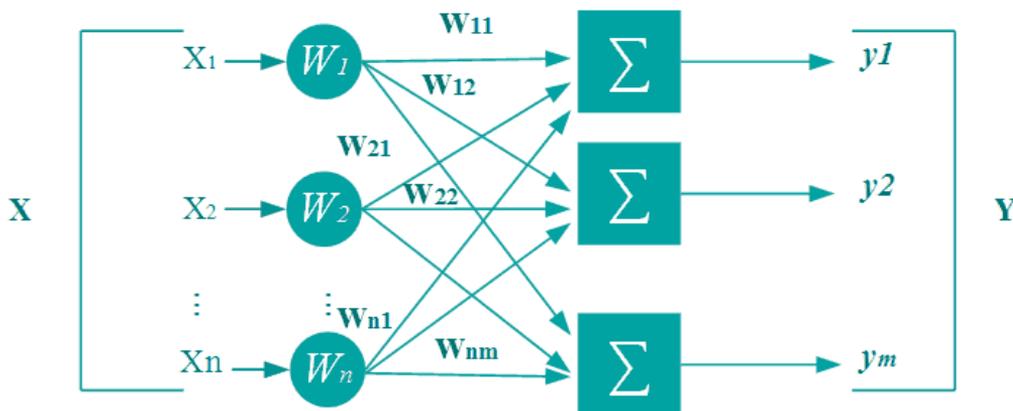
ИНС бывают нескольких видов:

- Однослойные
- Многослойные
- Сети прямого распределения

Однослойные нейронные сети

Однослойная нейронная сеть (Single-layer neural network) —в ней сигналы с входного слоя сразу подаются на выходной слой. Он производит необходимые вычисления, результаты которых сразу подаются на выходы.

Однослойная ИНС. Рис. 8.



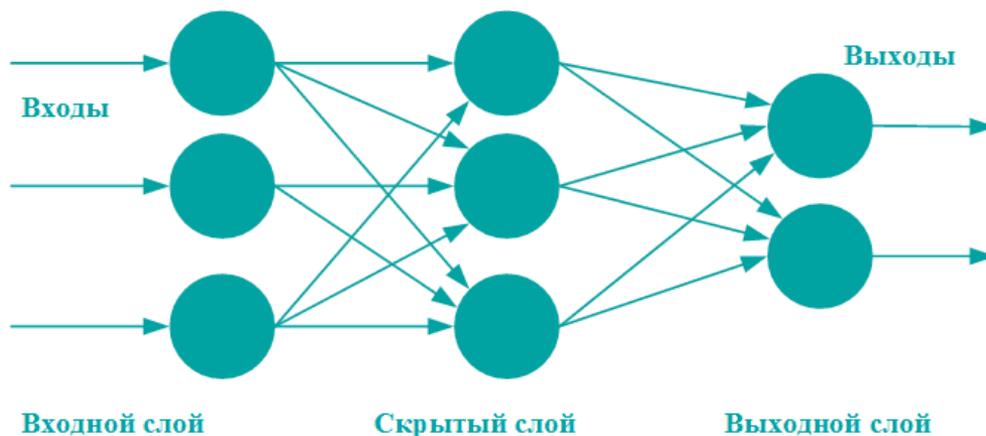
На схеме однослойной нейронной сети, сигналы x_1, x_2, \dots, x_n поступают на входной слой (который не считается за слой нейронной сети), а затем сигналы распределяются на выходной слой обычных нейронов. Нейроны соединены друг с другом стрелками. Над стрелками расположены веса соответствующих связей (весовые коэффициенты).

Многослойная нейронная сеть.

Многослойная нейронная сеть (англ. Multilayer neural network) — нейронная сеть, состоящая из входного, выходного и расположенного(ых) между ними одного (нескольких) скрытых слоев нейронов.

Помимо входного и выходного слоев эти нейронные сети содержат промежуточные, **скрытые слои**. Такие сети обладают гораздо большими возможностями, чем однослойные нейронные сети, однако методы обучения нейронов скрытого слоя были разработаны относительно недавно.

Многослойные нейронные сети. Рис. 9.



- Скрытые слои находятся между входными и выходными слоями
- Скрытые слои преобразуют входной сигнал в промежуточный
- Выходной сигнал собирается из промежуточных результатов, образованных в скрытых слоях

Вывод: Многослойная искусственная нейронная сеть имеет большие возможности по сравнению с однослойной, так как в скрытых слоях происходит преобразование входных сигналов в промежуточные результаты, которые собираются и передаются на выход. В этой ИНС может быть несколько скрытых слоев.

Сети прямого распространения

Сети прямого распространения (Feedforward neural network) (feedforward сети) — ИНС в которых сигналы идут строго слева направо от входного к выходному слою.

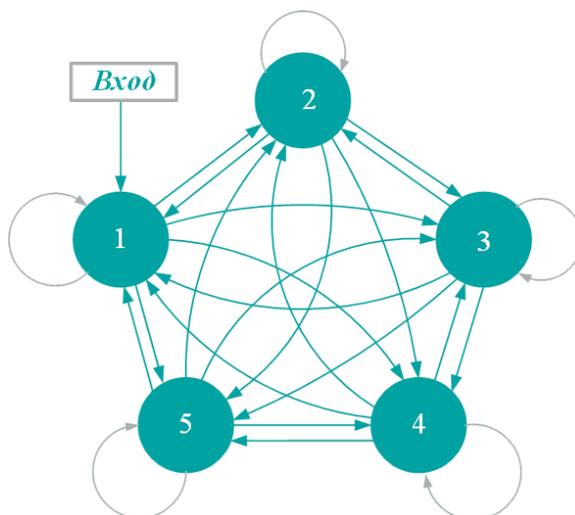
Эти сети широко используются под решение определенных задач: распознавание, кластеризация и прогнозирование.

Сети с обратными связями

Сети с обратными связями (Recurrent neural network) — искусственные нейронные сети, в которых сигнал подаваемый на выход, может вновь подаваться на вход.

В сетях такого типа сигнал может идти и в обратную сторону.

Сеть с обратными связями. Рис. 10.



ИНС с обратными связями имеют преимущество перед ИНС прямого распространения, так как выходы нейронов могут возвращаться на входы. Это означает, что выход какого-нибудь нейрона определяется не только его весами и входным сигналом, но еще и предыдущими выходами (так как они снова вернулись на входы). С помощью таких сетей можно создавать нейросети, восстанавливающие или дополняющие сигналы. Другими словами такие нейросети имеют свойства кратковременной памяти (как у человека).

Вопросы:

1. Назовите основные отличия видов ИНС
2. Приведите пример задачи, которую должна выполнить ИНС, какой вид ИНС вы используете?
3. Есть ли скрытые слои в однослойной ИНС?
4. Какую роль выполняют скрытые слои?
5. Можно ли сеть с обратными связями назвать циклическими?
6. Какую информацию несут нейроны возвращаясь на вход?
7. Какой вид ИНС имеет свойства кратковременной памяти (как у человека)?

§ 7.

Обучение ИНС

Словарь

- Обучение нейросети
- Обучение ИНС с учителем
- Обучение ИНС без учителя

Обучение нейросети

Если **ИНС** это связка искусственных нейронов, в которые на входе подается сигнал, внутри скрытых слоев происходит изменение параметров сигнала, а на выход подается некий результат, часто бессмысленный, значит нам необходимо влиять на параметры до тех пор, пока на выход мы получим нужный нам сигнал.

Параметры на которые мы можем повлиять и какой результат получим на выходе

Количество искусственных нейронов – при уменьшении или увеличении количества нейронов в ИНС не приведет к существенному изменению сигнала на выходе, так как при увеличении количества нейронов происходит увеличение вычислительных элементов и всю ИНС делает более увесистой и тяжелой, а сами вычислительные операции проходят одинаково.

Сумматор-сумматор выполняет одну строго заданную функцию-сложение. Если мы его заменим или уберем совсем, то сломаем сам ИН (Искусственный нейрон).

Функция активации - Если менять у каждого нейрона функцию активации, то мы получим слишком разношерстную и неконтролируемую нейронную сеть. К тому же, в большинстве случаев нейроны в нейронных сетях одного типа. То есть они все имеют одну и ту же функцию активации.

Остается только один вариант – менять веса связей.

Обучение нейронной сети (Training) — поиск такого набора весовых коэффициентов, при котором входной сигнал после прохода по сети преобразуется в нужный нам выходной.

Для того, чтобы нейросеть могла выдавать результат по большой базе данных, ей необходима способность обобщать какие-то признаки по всем вводным данным. С этой целью и создаются обучающие выборки.

Обучающая выборка (Training set) — конечный набор входных сигналов (иногда вместе с правильными выходными сигналами), по которым происходит обучение сети.

После обучения сети, то есть когда сеть выдает корректные результаты для всех входных сигналов из обучающей выборки, ее можно использовать на практике.

Для новой нейросети обязательно провести оценку качества ее работы. Такая оценка производится на тестовой выборке.

Тестовая выборка (Testing set) — конечный набор входных сигналов (иногда вместе с правильными выходными сигналами), по которым происходит оценка качества работы сети.

Вывод: «Обучение сети» это – подбор правильного набора весов, на тестовой выборке, для оценки качества работы.

Существуют два подхода для «Обучения нейросети», приводящие к разным результатам: обучение с учителем и обучение без учителя.

Обучение с учителем

Суть данного подхода заключается в том, что вы даете на вход сигнал, смотрите на ответ сети, а затем сравниваете его с уже готовым, правильным ответом.

Важный момент. Не путайте правильные ответы и известный алгоритм решения! Вы можете обвести пальцем лицо на фото (правильный ответ), но не сможете сказать, как это сделали (известный алгоритм). Тут такая же ситуация.

Затем, с помощью специальных алгоритмов, вы меняете веса связей нейронной сети и снова даете ей входной сигнал. Сравните ее ответ с правильным и повторяете этот процесс до тех пор, пока сеть не начнет отвечать с приемлемой точностью (как я говорил в 1 главе, однозначно точных ответов сеть давать не может).

Обучение с учителем (Supervised learning) — вид обучения сети, при котором ее веса меняются так, чтобы ответы сети минимально отличались от уже готовых правильных ответов. В роли учителя может выступить компьютерная программа.

Обучение без учителя

Обучение без учителя применяют тогда, когда у нас нет правильных ответов на входные сигналы. В этом случае вся обучающая выборка состоит из набора входных сигналов.

При таком обучении нейросеть начинает выделять общие признаки и характеристики, т.е. делает кластеризацию. Например, сеть получает на вход фотографии собак, кошек, хомяков, со временем сеть начинает выдавать сигналы трех типов, которые отвечают за объекты на входе.

Сюрприз состоит в том, что мы не знаем какие именно признаки и характеристики может выделить нейросеть при работе без учителя, с каждым новым поступлением данных на вход, сеть все более подробно кластеризирует данные. И очень часто мы можем на выходе получить неожиданный результат. Например, если нейросети дать самостоятельно на вход тысячи амбулаторных карт больных, и их лечение, то сеть может выделить общие признаки заболевания или более удачного варианта лечения, которые человек не заметил и не распознал.

Обучение без учителя (Unsupervised learning) — вид обучения сети, при котором сеть самостоятельно классифицирует входные сигналы. Правильные (эталонные) выходные сигналы не демонстрируются.

Вывод. В этой главе вы узнали все о структуре искусственного нейрона, а также получили полное представление о том, как он работает (и о его математической модели). Более того, вы теперь знаете о различных видах искусственных нейронных сетей: однослойные, многослойные, сети с обратными связями. Вы также ознакомились с тем, что представляет собой обучение сети с учителем и без учителя.

Вопросы:

1. Приведет ли к существенному изменению результата увеличение или уменьшение количества нейронов?
2. Какую функцию выполняет сумматор?
3. На что влияет изменение веса связей?
4. Что такое обучение нейронной сети?
5. Чем существенно отличаются обучение с учителем и обучение без учителя?
6. Что такое «Обучение сети»?

Итоговые вопросы

1. Что такое ИИ?
2. Что такое нейронная сеть?
3. Какие бывают нейроны?
4. Как выглядит математическая модель искусственного нейрона (формула)?
5. Опишите модель искусственного нейрона.
6. Какую роль играет вес поступающего входного сигнала?
7. Что такое сумматор, какую роль он выполняет?
8. Для чего нужна функция активации?
9. Какие виды ИНС бывают?
10. Какие задачи выполняют ИНС?
11. Все ли виды ИНС подходят для решения любой задачи?
12. Чем отличается обучение ИНС с учителем и без учителя, назовите существенную разницу.

Задача для закрепления.

На вход подается 4 сигнала:

2. 0,5
- 3. 2
1. 4
5. -1

Используя формулу расчета взвешанной суммы, определите вес сигнала, который нейрон подаст на выход?

Ответы.

§4. Задание 1.

1. 0 (0*1)
2. 1 (1*4)
3. 1 (1*1)
4. 0 (0*5)

Взвешанная сумма равна.

$$\text{net} = \sum_{i=1}^n x_i w_i = 0 + 4 + 1 + 0 = 5$$

так как порог $b=5$ по условию задачи, а взвешанная сумма должна быть \geq порога (b), следовательно, нейрон подаст на выход сигнал равный 1, и вывод о том, что пончик будет куплен.

Ответ: нейрон подаст на выход сигнал равный 1. Пончик будет куплен

§7. Задача для закрепления

$$\text{net} = \sum_{i=1}^n x_i w_i = 2 * 0,5 + (-3) * 2 + 1 * 4 + 5 * (-1) = -6$$

Ответ: Нейрон на выход согласно формуле расчета взвешанной суммы подаст $= -6$

Практикум

Разберем популярную задачу с Ирисами «Ирисы Фишера»

Ирисы Фишера—это набор данных для задачи классификации. Этот набор данных является классическим.

Задача.

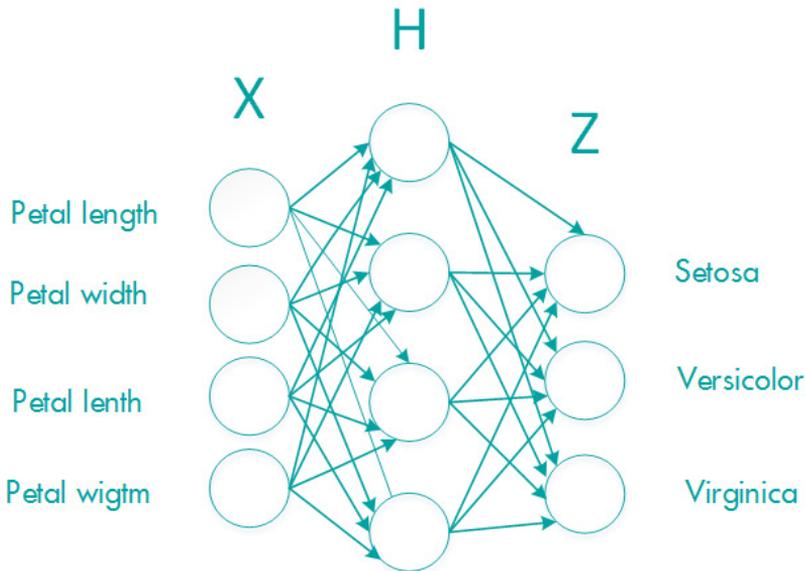
Ирисы Фишера состоят из данных о 150 экземплярах ириса, по 50 экземпляров из трёх видов:

- Ирис щетинистый (Iris setosa).
- Ирис виргинский (Iris virginica).
- Ирис разноцветный (Iris versicolor).

Для каждого экземпляра измерялись четыре характеристики (в сантиметрах):

- Длина чашелистника (sepal length).
- Ширина чашелистника (sepal width).
- Длина лепестка (petal length).
- Ширина лепестка (petal width).

На входе у нас есть 4 класса(характеристики) — X, также нам понадобится всего один внутренний слой — H, в нём будет 10 нейронов (выбирается методом подбора), далее на выходе мы имеем 3 класса, которые зависят от характеристики цветов — Z. Получается вот такая конструкция сети:



Распишем математическое обоснование для задачи.

Внутренний слой

$$t=x*W+b$$

$$h=relu(t)$$

Выходной слой

$$t=x*W+b$$

$$z = \text{softmax}(t)$$

Softmax-это функция, превращающая произвольный вектор, в набор вероятностей.

$$\text{softmax}(t) = \frac{e^{t_i}}{\sum_j e^{t_j}}$$

В качестве функции активации используем функцию Relu () по всем элементам t:

$$\text{Relu}(t) = \max(0, t)$$

Язык программирования Python.

Нам понадобится две библиотеки:

- Базовая библиотека языка Python, для работы с линейными данными, NumPy.
- Базовая библиотека языка Python, для «рандомизации» значений, random.

Для начала нам необходимо импортировать библиотеки numpy и random:

```
import numpy as np
import random as rd
```

Теперь пропишем некоторые гиперпараметры:

```
INPUT_DIM = 4 #кол-во входных нейронов
OUT_DIM = 3 #кол-во выходных нейронов
H_DIM = 10 #кол-во нейронов в скрытом слое
```

Теперь зададим входной вектор и его веса (вначале рандомим данные, для получения реальной картины весов):

```
x = np.random.randn(INPUT_DIM)
w1 = np.random.randn(INPUT_DIM, H_DIM)
b1 = np.random.randn(H_DIM)
w2 = np.random.randn(H_DIM, OUT_DIM)
b2 = np.random.randn(OUT_DIM)
```

Расписываем вложенный слой — наше математическое обоснование:

```
t1 = x @ w1 + b1
h1 = relu(t1)
```

Точно также сделаем и для остальных.

Теперь обернём наш код в функцию:

```
def predict(x):
    t1 = x @ W1 + b1
    h1 = relu(t1)
    t2 = h1 @ W2 + b2
```

```
z = softmax(t2)
print('z =', z)
return z
```

Оформим функцию relu():

```
def relu(t):
    print('relu:', np.maximum(t, 0))
    return np.maximum(t, 0)
```

Теперь добавим softmax():

```
def softmax(t):
    out = np.exp(t)
    print('softmax:', out / np.sum(out))
    return out / np.sum(out)
```

Добавим вызов функции predict(), также class_names — имена выходных классов и вывод результатов предсказания:

```
probs = predict(x)
pred_class = np.argmax(probs)
class_names = ['Setosa', 'Versicolor', 'Virginica']
print('Predicted:', class_names[pred_class])
```

Так как используем рандомные значения входных коэффициентов и весов, результат будет случайный.

Весь код сразу:

```
import numpy as np
import random as rd
INPUT_DIM = 4
OUT_DIM = 3
H_DIM = 10
x = []
for i in range(4):
    x.append(float(input()))
print(x)
# Рандомно вводим значения гиперпараметров:
w1 = np.random.randn(INPUT_DIM, H_DIM)
w2 = np.random.randn(H_DIM, OUT_DIM)
b1 = np.random.randn(H_DIM)
b2 = np.random.randn(OUT_DIM)
def relu(t):
    print('relu:1', np.maximum(t, 0))
    return np.maximum(t, 0)
def softmax(t):
```

```

    out = np.exp(t)
    print('softmax:', out / np.sum(out))
    return out / np.sum(out)
def predict(x):
    t1 = x @ w1 + b1
    h1 = relu(t1)
    t2 = h1 @ w2 + b2
    z = softmax(t2)
    print('z =1', z)
    return z
t1 = x @ w1 + b1
h1 = relu(t1)
probs = predict(x)
pred_class = np.argmax(probs)
class_names = ['Setosa', 'Versicolor', 'Virginica']
print('Predicted:', class_names[pred_class])

```

Добавим полученные после обучения веса и входные данные:

```

w1 = np.array([[ 0.33462099,  0.10068401,  0.20557238, -0.19043767,  0.40249301, -0.00925352,
 0.00628916,  0.74784975,  0.25069956, -0.09290041 ], [ 0.41689589,  0.93211640, -0.32300143, -0.13845456,
 0.58598293, -0.29140373, -0.28473491,  0.48021000, -0.32318306, -0.34146461 ], [-0.21927019, -0.76135162,
-0.11721704,  0.92123373,  0.19501658,  0.00904006,  1.03040632, -0.66867859, -0.01571104, -0.08372566 ],
[-0.67791724,  0.07044558, -0.40981071,  0.62098450, -0.33009159, -0.47352435,  0.09687051, -0.68724299,
 0.43823402, -0.26574543 ]])
b1 = np.array([-0.34133575, -0.24401602, -0.06262318, -0.30410971, -0.37097632,  0.02670964,
-0.51851308,  0.54665141,  0.20777536, -0.29905165 ])
w2 = np.array([[ 0.41186367,  0.15406952, -0.47391773 ], [ 0.79701137, -0.64672799, -0.06339983
], [-0.20137522, -0.07088810,  0.00212071 ], [-0.58743081, -0.17363843,  0.93769169 ], [ 0.33262125,
0.18999841, -0.14977653 ], [ 0.04450406,  0.26168097,  0.10104333 ], [-0.74384144,  0.33092591,  0.65464737
], [ 0.45764631,  0.48877246, -1.16928700 ], [-0.16020630, -0.12369116,  0.14171301 ], [ 0.26099978,
0.12834471,  0.20866959 ]])
b2 = np.array([-0.16286677,  0.06680119, -0.03563594 ])

```

Вся программа:

```

import numpy as np
import random as rd
INPUT_DIM = 4
OUT_DIM = 3
H_DIM = 10
x = []
for i in range(4):
    x.append(float(input()))
print(x)
w1 = np.array([[ 0.33462099,  0.10068401,  0.20557238, -0.19043767,  0.40249301, -0.00925352,
 0.00628916,  0.74784975,  0.25069956, -0.09290041 ], [ 0.41689589,  0.93211640, -0.32300143, -0.13845456,
 0.58598293, -0.29140373, -0.28473491,  0.48021000, -0.32318306, -0.34146461 ], [-0.21927019, -0.76135162,
-0.11721704,  0.92123373,  0.19501658,  0.00904006,  1.03040632, -0.66867859, -0.01571104, -0.08372566 ],
[-0.67791724,  0.07044558, -0.40981071,  0.62098450, -0.33009159, -0.47352435,  0.09687051, -0.68724299,
 0.43823402, -0.26574543 ]])

```

```

-0.11721704, 0.92123373, 0.19501658, 0.00904006, 1.03040632, -0.66867859, -0.01571104, -0.08372566 ],
[-0.67791724, 0.07044558, -0.40981071, 0.62098450, -0.33009159, -0.47352435, 0.09687051, -0.68724299,
0.43823402, -0.26574543 ]])
b1 = np.array([-0.34133575, -0.24401602, -0.06262318, -0.30410971, -0.37097632, 0.02670964,
-0.51851308, 0.54665141, 0.20777536, -0.29905165 ])
w2 = np.array([[ 0.41186367, 0.15406952, -0.47391773 ], [ 0.79701137, -0.64672799, -0.06339983
], [-0.20137522, -0.07088810, 0.00212071 ], [-0.58743081, -0.17363843, 0.93769169 ], [ 0.33262125,
0.18999841, -0.14977653 ], [ 0.04450406, 0.26168097, 0.10104333 ], [-0.74384144, 0.33092591, 0.65464737
], [ 0.45764631, 0.48877246, -1.16928700 ], [-0.16020630, -0.12369116, 0.14171301 ], [ 0.26099978,
0.12834471, 0.20866959 ]])
b2 = np.array([-0.16286677, 0.06680119, -0.03563594 ])
# x = np.random.randn(INPUT_DIM)
# w1 = np.random.randn(INPUT_DIM, H_DIM)
# b1 = np.random.randn(H_DIM)
# w2 = np.random.randn(H_DIM, OUT_DIM)
# b2 = np.random.randn(OUT_DIM)
def relu(t):
    print('relu:1', np.maximum(t, 0))
    return np.maximum(t, 0)
def softmax(t):
    out = np.exp(t)
    print('softmax:', out / np.sum(out))
    return out / np.sum(out)
def predict(x):
    t1 = x @ w1 + b1
    h1 = relu(t1)
    t2 = h1 @ w2 + b2
    z = softmax(t2)
    print('z =1', z)
    return z
t1 = x @ w1 + b1
h1 = relu(t1)
probs = predict(x)
pred_class = np.argmax(probs)
class_names = ['Setosa', 'Versicolor', 'Virginica']
print('Predicted:', class_names[pred_class])
Работа программы: (рандомные значения)
relu: [0.72976987 0. 0. 5.78924864 5.49372503 0. 6.55090738 1.69118863 1.85742283
0.]
softmax:
[3.19968193e-06 2.15422388e-02 9.784454562e-01]
z= [3.19968193e-06 2.15422388e-02 9.78754562e-01]
Predicted: Verginica

```

Заключение

В 1 части учебного пособия «Искусственные нейронные сети для школьников» «Знакомство с ИИ и нейросетями», мы познакомились с ИИ, узнали из чего состоит искусственный нейрон, Что такое входы, веса и сумматор. Разобрали как работает функция активации и какими способами можно обучать нейронную сеть. На практике научились вычислять, с помощью формулы расчета взвешанной суммы, какие сигналы на выход подает нейрон. Разобрали классический пример задачи «Ирисы Фишера» на языке программирования Python.

В следующей 2 части учебного пособия «Искусственные нейронные сети для школьников» «Обучение нейронных сетей, профессия нейропрограммист», мы познакомимся с персептронами, их классификацией. Научимся на практике обучать персептрон, разберем алгоритм обучения и напишем программу распознавания объектов на языке программирования Python.

Официальный сайт учебного пособия

«Искусственные нейронные сети для школьников» proneyroset.ru

Содержание

I Часть

Знакомство с ИИ и нейросетями.

§ 1. Знакомство с искусственным интеллектом (ИИ).	4
§ 2. Биологический нейрон и биологическая нейронная сеть.	8
§ 3. Искусственный нейрон.	10
§ 4. Функция активации. Функция единого скачка.	14
§ 5. Функция активации. Сигмоидальные функции.	16
§ 6. Виды ИНС (Искусственных нейронных сетей).	19
§ 7. Обучение ИНС.	22
Практикум	26
Заключение	32

Литература

Рышард Тадусевич, Барбара Боровик, Томаш Гончаж, Бартошц Леппер Элементарное введение в технологию Нейронных сетей с примерами программ // электронная книга

Тарик Рашид. Создаем нейронную сеть // электронная книга.

Полезные ссылки

Нейронные сети для начинающих. Решение задачи классификации Ирисов Фишера // статья [url: <https://habr.com/ru/company/ruvds/blog/679988/>]

Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных // электронный ресурс // веб. ресурс machinelearning.ru

Матвеева С.С. Искусственные нейронные сети для школьников // электронное учебное пособие // веб ресурс. proneyroset.ru

