**Ростислав Романов** @EmeraldSoft

Программист в области ИИ

**3,0**

карма

**18,5**

рейтинг



написать



Профиль

**2**

Публикации

**5**

Комментарии

вчера в 22:54

## Разное → Нейронные сети за 1 день

tutorial

Учебный процесс в IT, Исследования и прогнозы в IT

Всем привет. В этот раз мы попытаемся разобраться с нейронными сетями без



### Зачем они нужны?

Для того чтобы понять зачем нужны нейронные сети, нужно разобраться с тем

Искусственные нейронные сети — это совокупность искусственных нейронов,

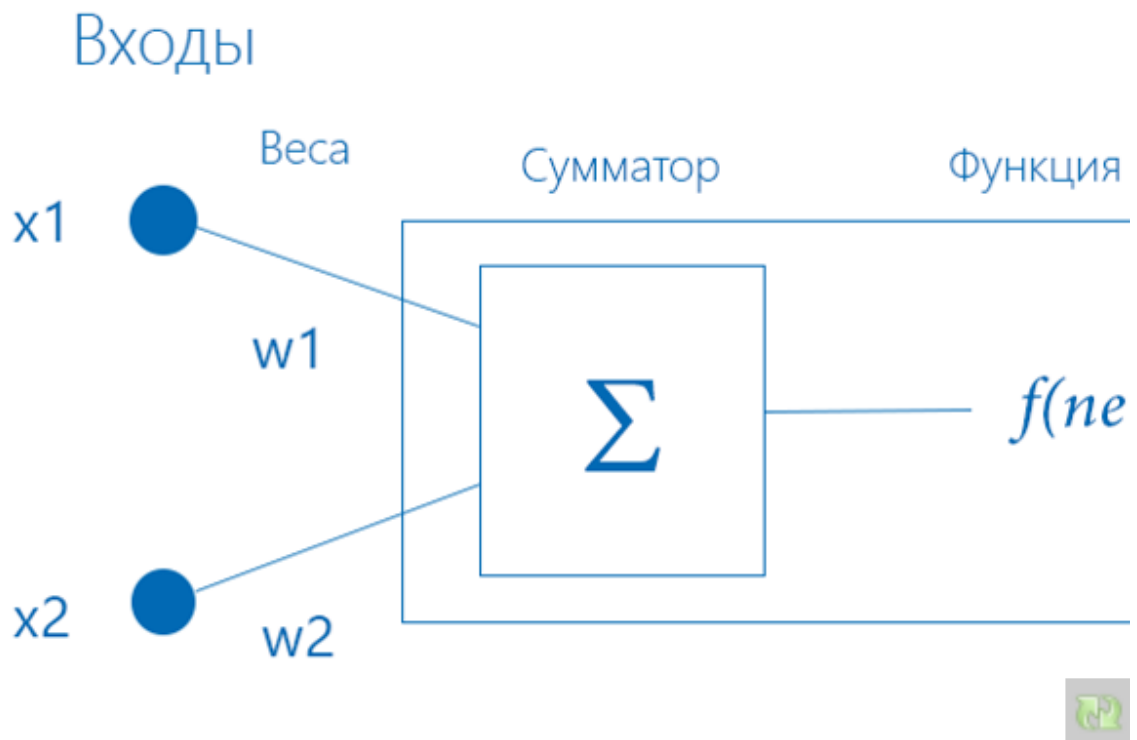
Искусственные нейронные сети нужны для решения сложных задач, например образов. Так же они применяются в области машинного обучения и искусственных в свои игры.

Главная особенность нейронных сетей — они способны обучаться.

## Искусственный нейрон

Перед тем как переходить к строению нейронных сетей, нужно разобраться с





За входы обозначены  $x_1$ ,  $x_2$ . На них поступают данные, либо в вещественном виде, либо в бинарном. Иногда приходится проводить нормализацию входных данных. Для этого достаточно: Количество входов зависит от задачи.

Так же мы имеем веса:  $w_1$ ,  $w_2$ . В них и заключается суть нейронных сетей, черпая из них информацию. Мы будем использовать их позднее.

Перед началом обработки данных, входы умножаются на соответствующие им веса. Результаты произведений поступают на нейрон и суммируются.  $(x_1 * w_1) + (x_2 * w_2)$

Например:  $x_1 = 1$ ;  $x_2 = 3$ ;  $w_1 = 0$ ;  $w_2 = -1$ ;

$$(1 * 0) + (3 * (-1)) = -3$$

Усвоим для себя, что количество весов должно соответствовать количеству вх

Далее результат суммирования поступает в блок нелинейного преобразования активации.

## Функция активации

Функции активации нужны для нормализации выходных данных. Допустим от хотели бы преобразовать его к 1 или 0.

Для такого имеется функция единого скачка. Когда в нейроне заряд превысил заряд ниже порога, то нейрон выдаст 0. Например,  $T$  — порог. Он равен 0. Ответ нейрона 0. Если заряд был бы больше, например равный 0.5 или 1 или 1. Порог устанавливается по вашему желанию, так как нейронная сеть всё равно

Когда нам необходимо преобразовать данные для большего выбора варианта, суммированный результат в вид от 0 до 1.

Для такого у нас имеется логистическая функция.

$$1/(1 + \exp(-a * net))$$



Или:

$$1/(1 + \exp(-net))$$

$a$  — степень крутизны на графике функции.

$net$  — результат суммирования

В нашем случае -3 преобразуется в 0.9525741268224334

В таком случае мы можем просклонять такой ответ к хорошему результату.

## Практика

Представим себе такую задачу. Мы хотели бы получить подсказку от нейронной сети: есть ли умная девушка или нет. Мы имеем такие входные данные:

Рост девушки, (в метрах)

На сколько она красива, (1% — 100%)

Есть ли ум, ( 1/0 )

Умеет ли готовить. (1/0)

Допустим установим все эти входы такими числами: 1.66, 100%, 1, 1

4 входа — 4 веса.  $t(\text{порог}) = 400$

Установим сейчас веса без обучения, как степень важности параметра:

1, 5, 5, 4.

Умножаем входы на соответствующие им по номеру веса:  $(1.66 * 1) + (100 * 5)$   
 $510.66 > 400$ , значит встречаться с девушкой можно, так как нейрон выдаст 1

Теперь попробуем проверить через логистическую функцию.  $1 / (1 + \exp(-510.66))$   
В данном случае мы тоже получили 1, значит встречаться можно. Но здесь у нас  
распределить ответы НС так:

$\text{net} \Rightarrow 0.80$ , встречаться можно.

$\text{net} < 0.80 \ \&\& \ \text{net} \neq 0.5$ , нужно ещё подумать.

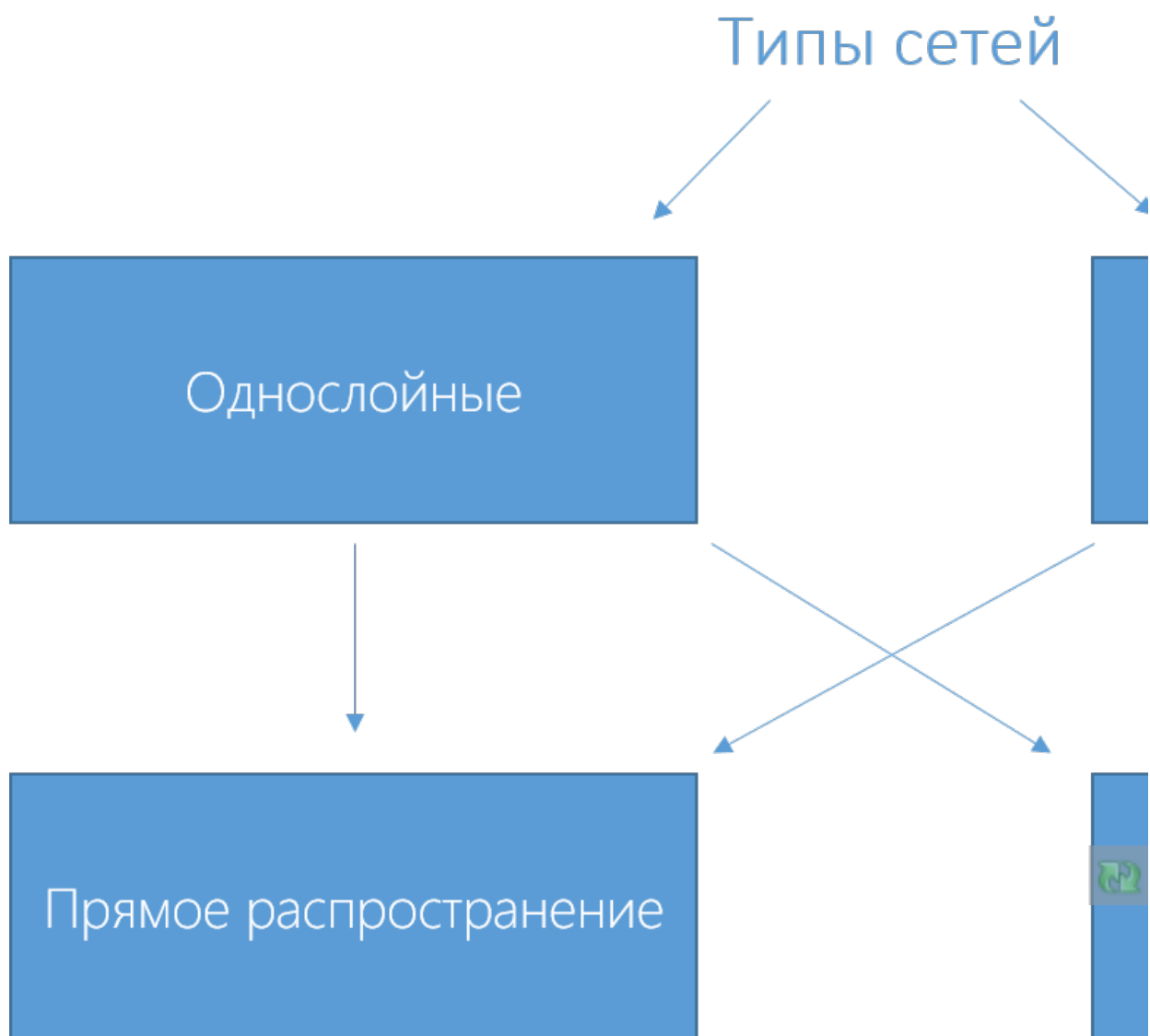
$\text{net} \leq 0.5$ , встречаться не нужно.

Надеюсь вам было понятно.

## Типы нейронных сетей

Теперь нужно понять строение нейронных сетей. Они разделяются на однослойные и обратными связями.

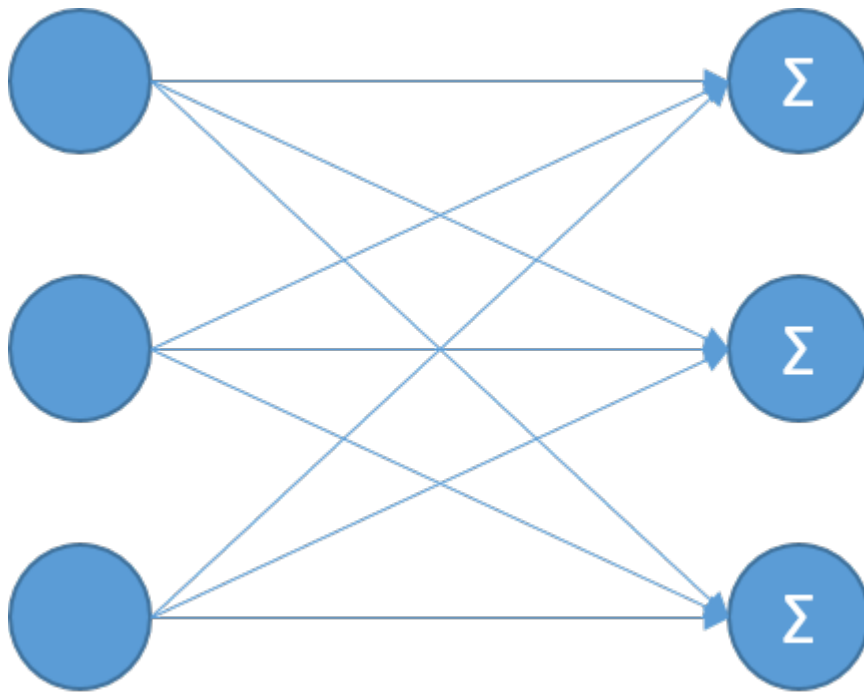




Единственное, что объединяет ОНС и МНС — это то, что входные нейроны не с  
обработку.

Однослойные НС

В таких НС данные с входных слоёв передаются сразу на выходные нейроны, и



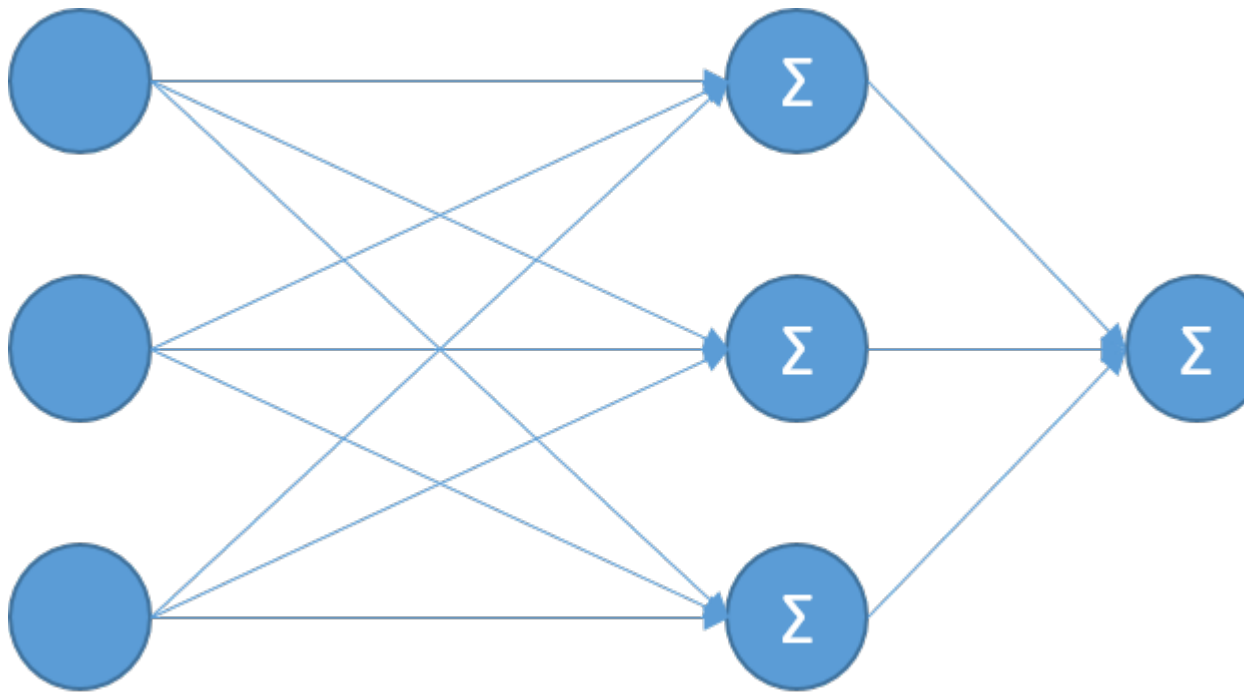
В данном примере три входных нейрона( $x_1, x_2, x_3$ ), три выходных нейрона( $out_1, out_2, out_3$ ). В таком случае, получается, три связи у каждого выходного нейрона.  
 $out_1: w_1, w_2, w_3$ .  $out_2: w_1, w_2, w_3$ .  $out_3: w_1, w_2, w_3$ .  
Каждая связь может иметь разное значение.

Как нейрон обрабатывает сигнал — я писал выше.



## Многослойные НС

Такие НС работают гораздо сложнее, но и способны они на большее.



Главная их отличительная особенность — они имеют скрытые(обрабатывающи  
 Названы они так, из-за того, что мы не видим какие сигналы они передают на  
 Сигналы они обрабатывают, часто имеют логистическую функцию активации,  
 количеством слоёв и нейронов в них вам нужно разобраться самим, ибо для эт

#### Сети с прямыми связями



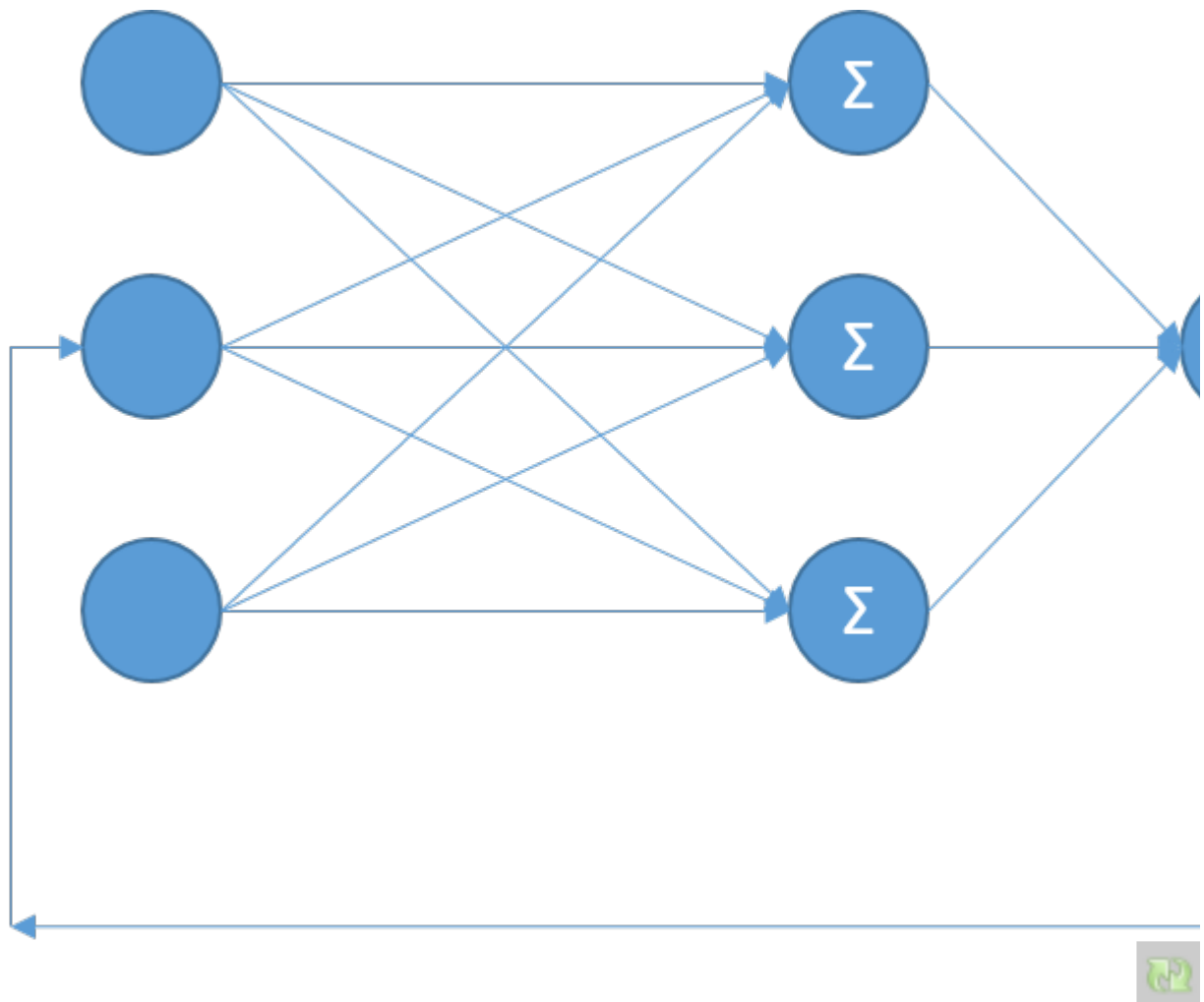
Сети с прямыми связями могут быть однослойные и многослойные, вы уже вид  
 связи из входов направляются к выходному слою или к скрытому слою и т.д к

Такие сети способны на прогнозирование, линейную интерполяцию, распознае  
 и многое другое.

#### Сети с обратными связями

Данные сети имеют как прямые связи, так и **обратные**.





Такие нейронные сети могут иметь разное количество слоёв, или могут быть в особенности — они имеют ассоциативную (кратковременную) память. На таких нейронных сетях работают чат-боты. Так же они могут применяться в управлении, решать задачи классификации.

## Перцептрон

Перцептрон — это нейронная сеть. Он имеет входной слой (S), слой обработки (A) и выходной слой (R).

Нейроны S слоя входные. Они могут находиться в состоянии возбуждения (1), или в состоянии покоя (0). Вы можете использовать и другие числа. Так же вы можете использовать и другие функции активации — ваше число.

Прежде чем попасть в A слой, сигналы с S слоя должны пройти по весам SA, значения которых находятся в диапазоне от -1 до 1 в вещественном виде. Далее, сигналы поступают в A слой, суммируются. Далее, сигналы, которые идут по AR весам. Здесь они могут иметь уже любые значения. В нейронах R слоя сигналы так же обрабатываются, проходят через функцию активации.

## Обучение

Вот мы и перешли к самой важной, но и в то же время очень сложной части. О котором изменяются весовые коэффициенты. Собственно, благодаря этим нейронные сети.

### Правило Хебба №1

Данное правило действует чаще всего с однослойными персептронами, входными и выходными числами у вас оно работать не будет.

Работает оно просто:

1. Если нейронная сеть выдаёт правильный результат, то весовые коэффициенты не изменяются.
2. Если нейронная сеть ошиблась и не дала правильный ответ(распознала не то), то весовые коэффициенты уменьшаются.
3. Если нейронная сеть ошиблась и отвергла правильный ответ, то весовые коэффициенты увеличиваются.

### Правило Хебба №2

Данное правило работает со всеми числами. По-другому оно называется delta rule, в которых нужно разобраться, так как delta правило присутствует в методах обучения нейронных сетей.



Для delta формулы нам необходимо знать ошибку сети. Чаще всего, ошибка — это разность правильного и неправильного ответов.

$$err = d - b$$

Где  $d$  — правильный ответ,  $b$  — ответ сети.

Теперь рассмотрим саму формулу:

$$w(t + 1) = w + err * n * x_i$$

$w(t + 1)$  — новый весовой коэффициент.

$w$  — старый весовой коэффициент

$err$  — ошибка(разность правильного ответа и ответа сети)

$n$  — скорость обучения

$x_i$  — значение которое пришло на  $i$ -ый вход

Я думаю, что вам не совсем понятна скорость обучения и как её искать. Ищется часто ставят 0.001, 0.0001 или 1.

## Метод обратного распространения ошибки

Это самый лучший метод обучения сетей, который используется как с многослойными, так и с однослойными. Чем-то он может напомнить вам  $\delta$  правило, ибо формула та же.

В данном случае, нам необходимо создать счётчик ошибок. Если показатель счётчика достигнет заданного порога, то обучение останавливается. Конечно же, в данном случае нам необходимо иметь обучающую подвыборку.

Алгоритм такого обучения следующий:

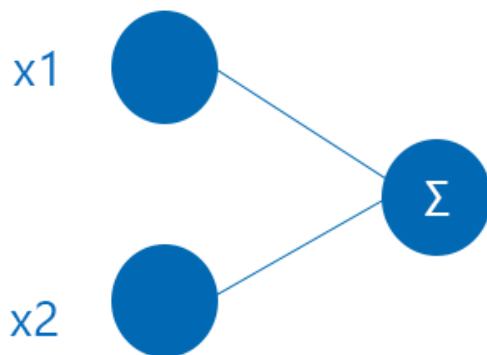
1. Инициализация весов случайными значениями.
2. Выбрать обучающие данные, подать на входы сети.
3. Вычислить выход сети.
4. Вычислить разность между правильным ответом и неправильным.
5. Корректируем веса для минимизации ошибки.
6. Повторяем 2 и 5 шаг, пока ошибка не достигнет 0 или приемлемого уровня.

## Практика

Вот теперь мы практикуем свои знания. Сейчас мы попробуем сделать однослойную сеть, которая будет работать с методом обратного распространения ошибки.



Пусть он строит логическую таблицу функций 'AND' и 'OR'. Кто знаком с алгеброй булевых функций?



Начнём описывать эту сеть через ООП. Мы видим, что у нас два входных нейрона

```
public class NeuralNetwork {  
    static double enters[] = new double[2]; // создаём входы  
    static double out; // храним выход сети  
    static double[] weights = {0,0}; // весовые коэффициенты
```

Теперь нам необходимо создать матрицу готовых ответов и входов для таблиц

```
static double tableOfLearn[][] = {  
    {0,0,0},  
    {1,0,0},  
    {0,1,0},  
    {1,1,1}  
};
```

Создаём обработку входных данных. Здесь не так всё сложно. Если вы помните соответствующие им по индексу веса и результаты произведений суммируются. Если выход  $> 0.1$ , то НС даёт 1.

```
public static void summator(){
```

```

        out = 0; обнуляем выход
        for ( int i = 0; i < enters.length; i++ )
            out+=enters[i]*weights[i]; // вход * вес, суммируем.

        if ( out > 0.1 ) out=1; else out=0; // функция активации
    }

```

Теперь пишем обучение сети. Метод обратного распространения ошибки. Создаём копию входы НС входные данные из обучающей таблицы. Запускаем обратное распространение ошибки, используем delta правило.

```

public static void train(){
    double gError = 0; // создаём счётчик ошибок
    int it = 0; // количество итераций
    do {
        gError = 0; // обнуляем счётчик
        it++; // увеличиваем на 1 итерации
        for ( int i = 0; i < tableOfLearn.length; i++ ){
            enters = java.util.Arrays.copyOf(tableOfLearn[i],
                tableOfLearn[i].length - 1); // копируем в v
            summator(); // суммируем
            double error = tableOfLearn[i][2] - out; // получаем
            gError+=Math.abs(error); // суммируем ошибку в модуль
            for ( int j = 0; j < enters.length; j++ )
                weights[j]+=0.1*error*enters[j]; // старый в
        }
    } while(gError!=0); // пока gError не равно 0, выполняем код
}

```

Теперь мы запускаем тестер НС. Включаем обучение; НС обучилась; проверяем

```

train();
for ( int p = 0; p < tableOfLearn.length; p++ ){
    enters = java.util.Arrays.copyOf(tableOfLearn[p],
        tableOfLearn[p].length - 1);

    summator();

    System.out.println(out);
}

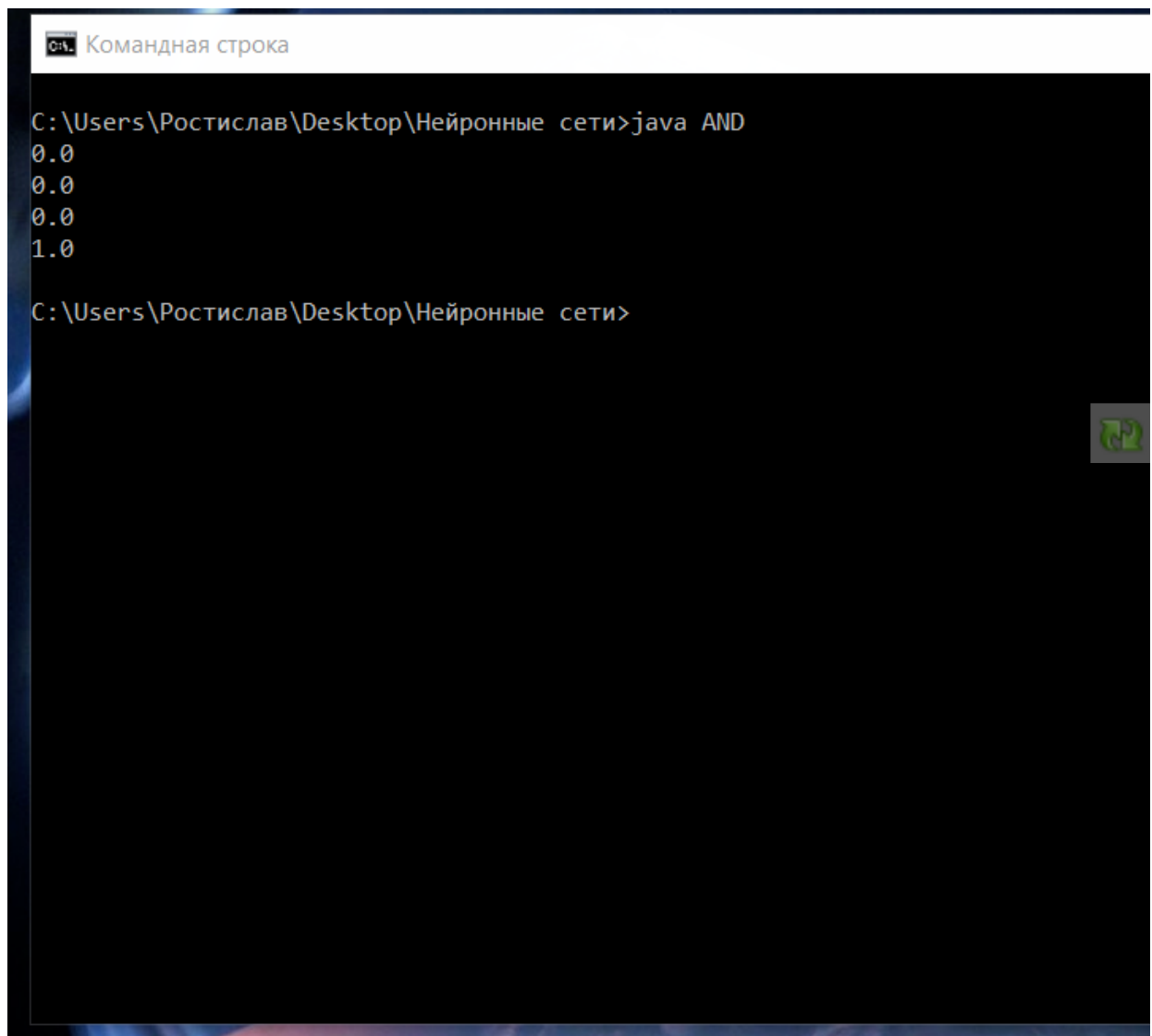
```

```
}
```

Включаем тестирование при запуске программы.

```
public static void main(String[] args) {  
    new NeuralNetwork().test();  
}
```

Вот такой результат мы получаем:



```
Командная строка  
C:\Users\Ростислав\Desktop\Нейронные сети>java AND  
0.0  
0.0  
0.0  
1.0  
C:\Users\Ростислав\Desktop\Нейронные сети>
```

## Заключение

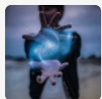
Конечно, многое может быть вам было не понятно. На сайте я буду писать мнс проектов с НС, поэтому объяснять о них я буду не один раз. Я надеюсь, что ко

внимание, удачного кода и послушного ИИ.

neural networks, java, ИИ, artificial intelligence



👁 14k ★ 404



**Ростислав Романов @EmeraldSoft**

Программист в области ИИ

карма рейтинг

3,0 18,5

## Похожие публикации

**+16** DetectNet: Deep Neural Network для Object Detection в DIGITS

👁 6,3k ★ 61 💬 12

**+16** Deconvolutional Neural Network

👁 41k ★ 225 💬 0

**+30** Алгоритм Self-Organizing Incremental Neural Network (SOINN)

👁 17,8k ★ 204 💬 7



## Самое читаемое

Сейчас

Сутки

Неделя

Месяц

**+29** Нейронные сети за 1 день

👁 14k ★ 404 💬 46

+5 Петербургский ИТ-экспорт недооценили на крупную сумму

👁 705 ★ 1 💬 0

+38 Как сделать девушке предложение при помощи социальной инженерии

👁 15,4k ★ 73 💬 15

+28 Зарплаты «топов» ИТ-компаний вновь попали в топ

👁 15,3k ★ 28 💬 11

+19 5 практических советов по эксплуатации литий-ионных аккумуляторов

👁 682k ★ 312 💬 29

## Комментарии (46)

отслеживать новые: ☐ в почте ☒ в треке

 **inkер** 30 января 2017 в 23:14 # ★

В этом месяце ещё не было.

[ответить](#)



 **noonv** 31 января 2017 в 11:44 # ★ ↵ ↑

Но автор, всё-таки, в предпоследний день успел :)


[ответить](#)

 **lair** 30 января 2017 в 23:39 # ★

Начнём описывать это сеть через ООП.

Зачем?

[ответить](#)

 **EmeraldSoft** 30 января 2017 в 23:45 # ★ ↵ ↑

Просто, почему бы и нет?

[ответить](#)

 **lair** 30 января 2017 в 23:47 # ★ ↵ ↑

Потому что это неэффективно в практическом применении.

(не говоря о том, что это у вас, прямо скажем, весьма ложненькое ООП)

[ответить](#)





**EmeraldSoft** 30 января 2017 в 23:56 (комментарий был изменён)

# ★ h ↑

Если людям не нравится — они обычно меняют.

[ответить](#)



**lair** 30 января 2017 в 23:58 # ★ h ↑

Для этого надо понимать, как и куда менять. А для этого надо понимать, что в

[ответить](#)



**Coffin** 31 января 2017 в 10:17 # ★ h ↑

Предложите свое видение без ООП :) Всем будет интересно.

[ответить](#)



**lora** 31 января 2017 в 10:51 # ★ h ↑

Слово «матрицы» вам о чем-нибудь говорит?

[ответить](#)



**Coffin** 31 января 2017 в 10:53 # ★ h ↑

Следуй за белым кроликом?

Синяя или красная?

Нео

Тринити

или вы про другие?

[ответить](#)



**lora** 31 января 2017 в 10:57 # ★ h ↑

Про математические.

[ответить](#)



**Coffin** 31 января 2017 в 10:58 # ★ h ↑

Именно о них и думаю, когда говорят слово матрицы.

Но какое это имеет отношение к моему комментарию о том, что

[ответить](#)



**lora** 31 января 2017 в 11:15 # ★ h ↑

Но какое это имеет отношение к моему комментарию о том

Прямое. Но если вы задаете такие вопросы, если у вас они и  
весьма смутное представление о нейронных сетях.

[ответить](#)



**Coffin** 31 января 2017 в 11:20 # ★ h ↑

Если вы не заметили или не хотели заметить или еще как  
Я лишь предложил автору коммента про «зачем ООП» пр  
он себе видит код.

Просто вы не умеет читать то, что пишут люди, а читаете  
«аудит весов» вашей нейронной сети серого вещества.

[ответить](#)



**Iora** 31 января 2017 в 11:23 (комментарий был изменён) :

Вы молодец, всегда можете оправдать свое невежество

[ответить](#)



**Coffin** 31 января 2017 в 11:29 # ★ h ↑

Спасибо.

Я указал вам правильный путь, следуйте ему и все буд

[ответить](#)



**Iora** 31 января 2017 в 11:36 # ★ h ↑

С моей стороны остается только пожелать вам изучить  
возможно, вы осознаете, что есть иные методологии к  
реализации ИНС.

[ответить](#)



**Coffin** 31 января 2017 в 11:38 # ★ h ↑

Я еще раз вам советую читать то, что пишут, а не то, чт  
Я ни разу не сказал, что для ИНС надо писать только на

[ответить](#)



**Iair** 31 января 2017 в 11:54 (комментарий был изменён) # ★ h ↑

Как-то так:

```
% input layer
% adding bias
a1 = [ones(m, 1) X];

% hidden layer
z2 = a1 * Theta1';
a2 = sigmoid(z2);
% adding bias
a2 = [ones(size(z2, 1), 1) a2];
```

```
% output layer
z3 = a2 * Theta2';
a3 = sigmoid(z3);

% prediction
[a, p] = max(a3');
p = p';
```

(a — входы, z — результаты суммирования, p — выход)

Ну и да, это не какое-то "мое видение", это задачка по учебнику.

[ОТВЕТИТЬ](#)



**masai** 31 января 2017 в 19:01 # ★ h ↑

Обозначения как в курсе от Эндрю Ына. Угадал?

[ОТВЕТИТЬ](#)



**lair** 31 января 2017 в 19:04 # ★ h ↑

Да. Говорю же — по учебнику, ничего оригинального.

[ОТВЕТИТЬ](#)



**lair** 30 января 2017 в 23:57 # ★

Мне, кстати, отдельно интересно, чем приведенная вами сеть функционально отличается

[ОТВЕТИТЬ](#)



**Dimchansky** 31 января 2017 в 14:53 # ★ h ↑

А разве есть разница между нейронной сетью с одним нейроном и логистической регрессией?

[ОТВЕТИТЬ](#)



**lair** 31 января 2017 в 14:55 # ★ h ↑

Вот и мне кажется, что нет, но может я чего-то не понимаю?

[ОТВЕТИТЬ](#)



**FedyaShlyapkin** 31 января 2017 в 03:26 # ★

так что там с девушкой? нашли её с помощью нейросетей?

[ОТВЕТИТЬ](#)



**Zaicheg** 31 января 2017 в 07:22 # ★ h ↑

Тятя, тятя, наши сети  
Притащили мертвеца!

[ОТВЕТИТЬ](#)



**JDBiber** 31 января 2017 в 12:26 # ★ h ↑

Наверняка. Любая выше 400 метров ростом — подходящий кандидат!

[ответить](#)



**ProgrammerMicrosoft** 31 января 2017 в 04:44 # ★

Автору огромное спасибо за статью :)

[ответить](#)



**polybook** 31 января 2017 в 08:56 # ★ h ↑

Присоединяюсь к благодарности и жду продолжения.

[ответить](#)



**theWaR\_13** 31 января 2017 в 09:01 # ★ h ↑

Нда... Кажется только на хабре могут заминусовать сообщение с благодарностью авто

[ответить](#)



**GeMir** 31 января 2017 в 08:54 # ★

О LaTeX автор не слышал?

«Для этого достаточно:  $1/\text{number}$ » —?

«выходной слой®» — ??



В целом очень похоже на школьный реферат по теме из категории «мне бы на троечку».

[ответить](#)



**bromzh** 31 января 2017 в 09:15 # ★

Начнём описывать эту сеть через ООП.  
static, static, static

Нет автор, это не ООП, а обычный процедурный стиль.

[ответить](#)



**rkfg** 31 января 2017 в 09:58 # ★

Мне было бы более интересно почитать про практическое применение нейросетей, но без мало. Я нашёл хорошее объяснение с примерами и кодом на [Deeplearning4j](#), да и сама би как и сам язык Java, она ориентирована на инженеров, а не на учёных, позволяет решать Освещены свёрточные и рекуррентные сети, есть объяснения с картинками, а не со «стра этого маловато для полномасштабного применения, и хотелось бы в таком же стиле прод

Например, я довольно долго пытался обучить чат-бота на LSTM по аналогии [разработки](#) ( сложными, всё влезает на 8 Гб видеопамети, обучающих данных достаточно (чатлоги кони из наиболее часто встречающихся 10 тыс. слов и пунктуации, далее слова заменялись на одного и того же ника, и в качестве предсказываемого результата подавал следующий блс

а вторым брался следующий блок. Диалог, понятно, далеко не всегда получался, но ничег модели вёл себя не так, как должен, независимо от learning rate — болтался возле одного Тестовая выдача тоже не сильно коррелировала с ожидаемой и выглядела плохо на фоне несколько улучшил ситуацию, как и снижение размера минибатча, но в целом, желаемого

Потом был ещё эксперимент с увеличением размера выборки обучения, но там вылезли с объяснений, что это значит и как с этим бороться, я не нашёл. Проблема могла быть как в маскировании неиспользуемых элементов, так и в неверно выбранном подходе в принцип посимвольное предсказание с очень большой обучающей выборкой, по 1000 символов на выборки были по слов 20-30 максимум, а чаще намного меньше. В примере гугла и того м влияет на способность сети к обучению? Мне неизвестно.

Не отрицаю, что возможно, после изучения всей линейной алгебры, лежащей в основе этс намного лучше. Но я заметил, что в целом даже у специалистов подходы практикуются до обнаружили какие-то свойства, которые не всегда получается объяснить (например, пере наперёд даёт более устойчивую связь с предсказываемым). А современные библиотеки и абстрактный уровень, чтобы разработчику необязательно было досконально знать, как та какие бывают типы сетей, какие у них свойства, как прикинуть размер скрытых слоёв в за ширине, так и по времени, кстати, тоже не особо освещённая тема — в примерах ставитс никак не мотивируется), как выбрать learning rate, minibatch size, как выбрать функцию акт для конкретной задачи (и какими практическими свойствами они все отличаются, по форм применимы), как распараллелить работу на кластер, если он есть и т.п.

Вот про это всё в научно-популярном, нематематическом стиле, с картинками и примерам программирования я бы почитал с превеликим удовольствием. А про биологические анал какое отношение имеют писано-переписано уже везде. Равно как и полно материала «для двигаться.

[ответить](#)



**Coffin** 31 января 2017 в 10:22 # ★ h ↑

Думаю проблема обучения чат-ботов в том, что все ожидают что-то вроде:

Реплика 1

Ответ на реплику 1

Реплика 2

Ответ на реплику 2

А в реальной же жизни это

Реплика 1

Ответа на реплику 2

Реплика 2

продолжение ответа на реплику 1

Ответ на реплику 1 и реплику 2

Реплика 3 и 4

Ответ на 4 и 1

Ответ на 1 и 2

Ответ на 3 и 4

[ответить](#)



**rkfg** 31 января 2017 в 10:26 # ★ h ↑

Да, всё так, на обычных логах построить внятную и логичную беседу не получится. Г частично релевантного смешного бреда, всё же делалось это для развлечения. Мен выдавал хотя бы реплику из обучающей выборки в ответ на предыдущую. Но не пол обучать сильно дольше, я оставлял на несколько часов, но визуально в Web UI сред уверен, что был смысл продолжать. С нейросетями вообще ни в чём нельзя быть ув

[ответить](#)



**Akon32** 31 января 2017 в 10:37 # ★ h ↑

я довольно долго пытался обучить чат-бота

Чтобы бот сносно вёл диалоги, ему нужна какая-то модель мира. Если вы пытаетесь о модель должна появиться в нейросети в процессе обучения. Вероятно, в чат-логах нед модели (хотя я не знаю, может гигабайтов или терабайтов разноплановых чат-логов бу,

Я объединил реплики, идущие подряд от одного и того же ника, и в качестве предска реплик, после чего второй становился первым, а вторым брался следующий блок.

Насколько я понял по описанию, объём контекста беседы легко мог быть выбран непре так что нейросеть не могла сделать никаких выводов о том, что следует говорить.

[ответить](#)



**rkfg** 31 января 2017 в 11:13 # ★ h ↑



Меня бы устроил оверфиттинг для начала, пусть бы выдавала фиксированные уже вход. А так да, объём данных конечно же имеет значение, но и информации по оцен количества label'ов тоже не встречал. Понятно, что чем больше, тем лучше, и если ч приведёт только к оверфиттингу и вообще плохо, но сколько будет достаточно, скаж

Кроме того, я мог неверно воспринимать понятие score. Я знаю, что это не evaluation выборке, а скорее взвешенная разность между выходом сети и желаемым значение. Примеры в deeplearning4j показывают сходимость score к нулю, у меня он болтался. Learning rate тоже подбирается с трудом, скажем, на  $1e-3$  он может через пару итера бесконечность, а на  $1e-4$  вообще не оказывать заметного влияния. Зато через 200 и там. Как это интерпретировать, мне тоже не очень ясно.

[ответить](#)



**FedyuninV** 31 января 2017 в 15:11 # ★ h ↑

А что за пример в переворачиванием? Не осталось ли ссылки на статью с этим?

P.S. Тот самый случай, когда один комментарий гораздо лучше всей статьи.

[ответить](#)



**rkfg** 31 января 2017 в 15:25 # ★ h ↑

[Вот тут](#) неоднократно упоминается, что переворачивание текста привело к улучшени

в первых словах содержится больше всего смысла, и градиент от них должен идти в максимальную силу. Рекуррентные сети обучаются не только в «пространстве», но и последовательности), поэтому, очевидно, самые «старые» элементы будут иметь на перезаписаны новыми коррекциями весов от более поздних элементов последовательности переместить в конец, они окажут наибольшее значение при генерации ответа сети. И сработает не для всех языков — в японском наиболее значимые слова часто идут в начале предложения, оказывая влияние на культуру народа, когда прямолинейность и быстрый переход к делу считается походить вокруг да около. Но это мои догадки вилами по воде, конечно.

[ОТВЕТИТЬ](#)



**mwizard** 31 января 2017 в 10:59 # ★

И ни слова про bias. Разбирайтесь лучше с базовой структурой сетей, автор.

[ОТВЕТИТЬ](#)



**aso** 31 января 2017 в 11:08 # ★

2. Если нейронная сеть ошиблась и не дала правильный ответ(распознала неверно), то

«Ложное срабатывание», ака «ошибка второго рода»?

3. Если нейронная сеть ошиблась и отвергла правильный ответ, то весовые коэффициенты



Пропуск цели, ошибка первого рода.

С подсчётом слоёв не всё так однозначно — в «входной слой» не все не всегда учит нейроны, а непосредственно входные датчики.

С другой стороны — [сеть Хопфилда](#) — имеет «просто» один слой, без входного — зато св

[ОТВЕТИТЬ](#)



**mopsicus** 31 января 2017 в 11:14 # ★

Статья [полугодовой давности](#), тоже простыми словами, тоже про девушек, но более понятная

[ОТВЕТИТЬ](#)



**alex87is** 31 января 2017 в 12:55 # ★

Статья понравилась :) Спасибо. Ясно-понятно. Не без небольших грехов, но люблю когда оформляют и тп. Давно хотел прочитать такую, но все ленился))) А тут на работе выпал свободный вечер!

Чем больше читаю про Машинное обучение, тем больше понимаю, что самое «крутое» в :

[ОТВЕТИТЬ](#)



**Landgraph** 31 января 2017 в 14:31 # ★

Гм... Автор, спасибо за статью! Я так понимаю, что с того момента, как я последний раз делал школьные годы (начало нулевых) ничего принципиально не изменилось?

В то время всё сильно буксовало из-за недостаточной вычислительной мощности, конкретнее распараллеливания вычислений дабы повысить параллелизм работы этих «нейронов»...

А я-то думал... Но с другой стороны это и хорошо, можно смело «возвращаться» =)

[ответить](#)



**rkfg** 31 января 2017 в 15:11 # ★ h ↑

Случился синергетический эффект. С одной стороны, выросли вычислительные мощные нейросети на порядок в сравнении с топовым процессором, есть практически в каждом CUDA-кластерами. С другой стороны, интернет, который раньше был сильно меньше и корпусами данных на любой вкус, будь то книги, субтитры к фильмам, твиты или ревью. И конечно, после первого бума и забвения нейросетей появились оригинальные алгоритмы сетей (постарше). Их комбинация позволяет получать, например, описание изображений.

[ответить](#)



**masai** 31 января 2017 в 18:56 # ★ h ↑

Вообще, изменилось много чего. Просто большинство таких статей опирается на те же однослойные перцептроны, не используются. А значит и правило Хебба становится неактуальным. Не использовать методы оптимизации посерьезнее: Adam, RMSProp и так далее. Они популярны для начинающих, любят писать про сигмоиду, но реально гораздо лучше использовать обычные методы оптимизации, уже не справляющиеся, поэтому вводят одинаковые веса для чего делают на сверточных сетях. Вместо многослойных сетей используют довольно много внимания уделяют проблеме переобученности, вводят dropout и другие подобные.

В общем, очень много чего изменилось. Не судите о современном состоянии дел по та

[ответить](#)



**Timoschenko**

Большое спасибо!

Мне нравится как вы все разжевываете. Я не математик, но это интересно и мне тоже. И в математике трудно понять.

А у вас все понятно расписано.

Жду продолжение!











Ваш комментарий отправлен на модерацию

Комментарий будет виден другим пользователям в случае, если он будет одобрен администратором ресурса.



Продолжить

## Интересные публикации

-  Петербургский ИТ-экспорт недооценили на крупную сумму  0
-  Как мы Redis Cluster готовили  0
-  В Сан-Франциско открылась полностью автоматизированная кофешкофа  0
-  Azure понятным языком [Шпаргалка]  0
-  Логика сознания. Часть 10. Задача обобщения  0

