什么是人工神经网络？

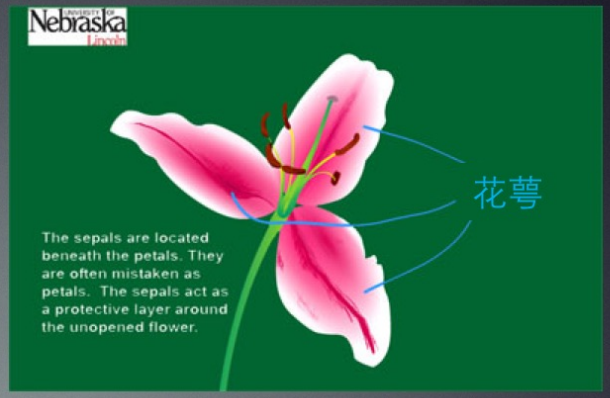
鸢尾花分类

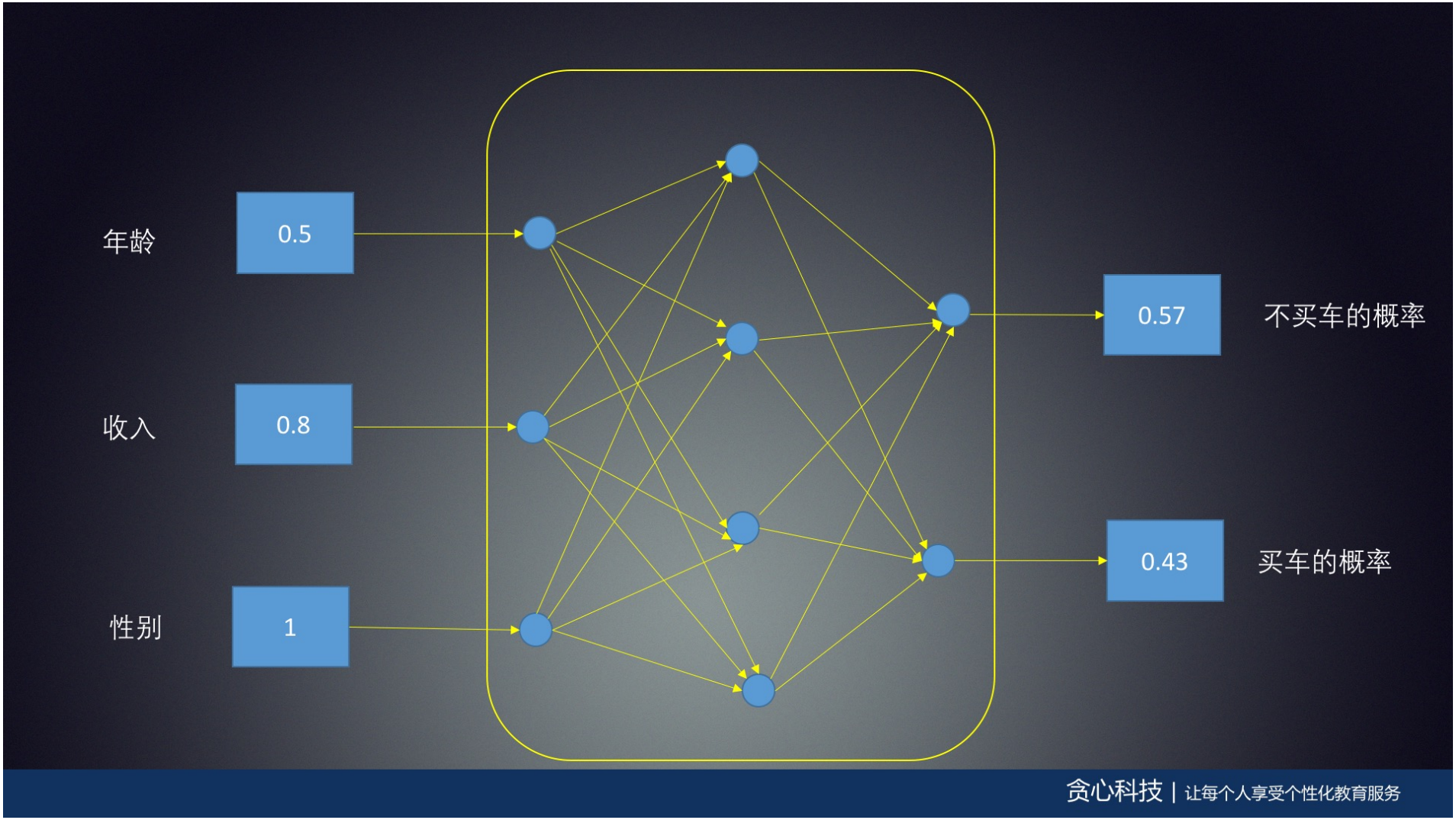
山鸢尾花(Iris Setosa)

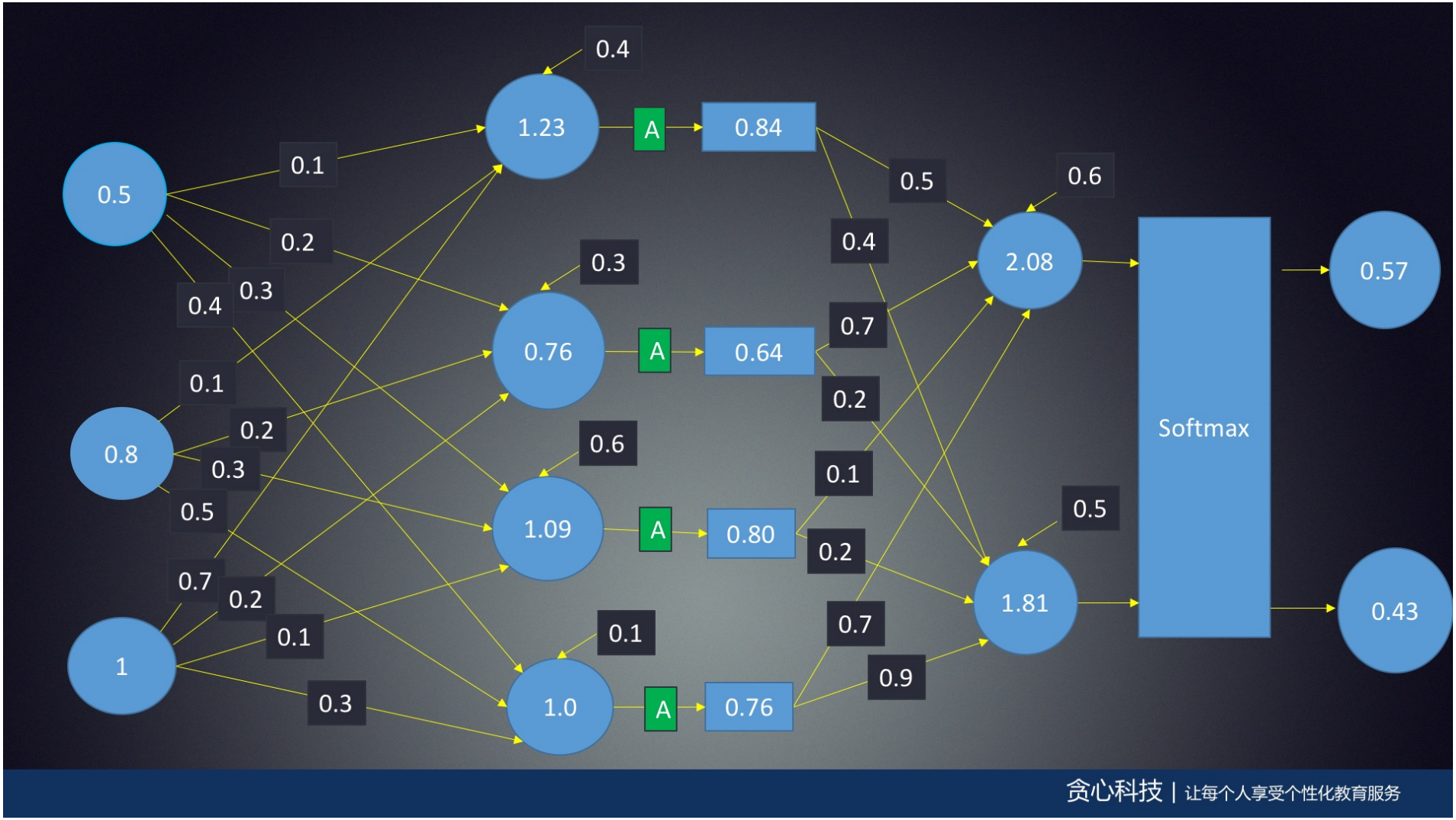
变色鸢尾花
(Iris-
Versicolor)

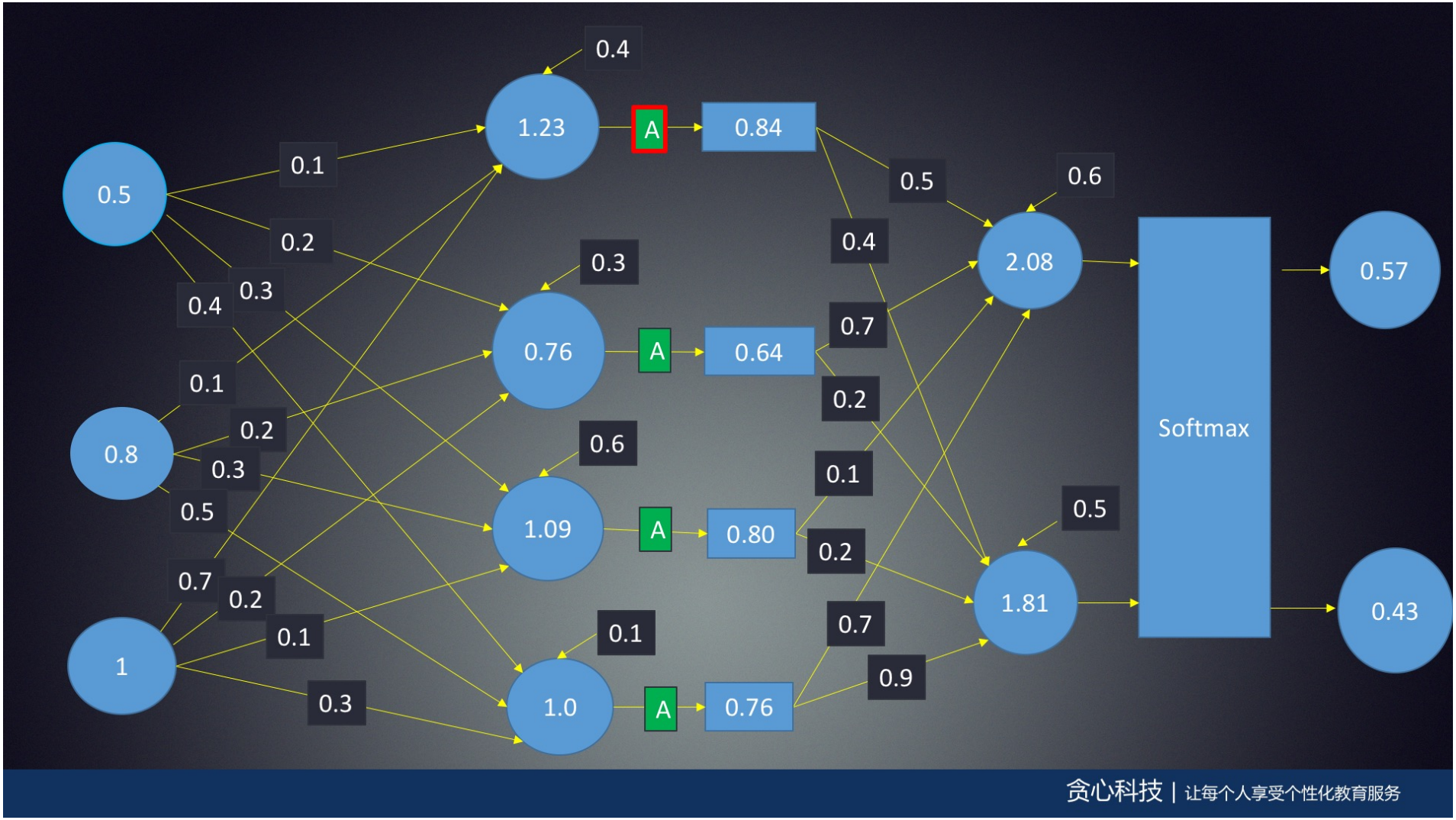
维吉尼亚鸢尾花
(Iris Virginica)


Nebraska
Lincoln
花瓣
The petals act as a colorful display to atract potential pollinators to the flower.

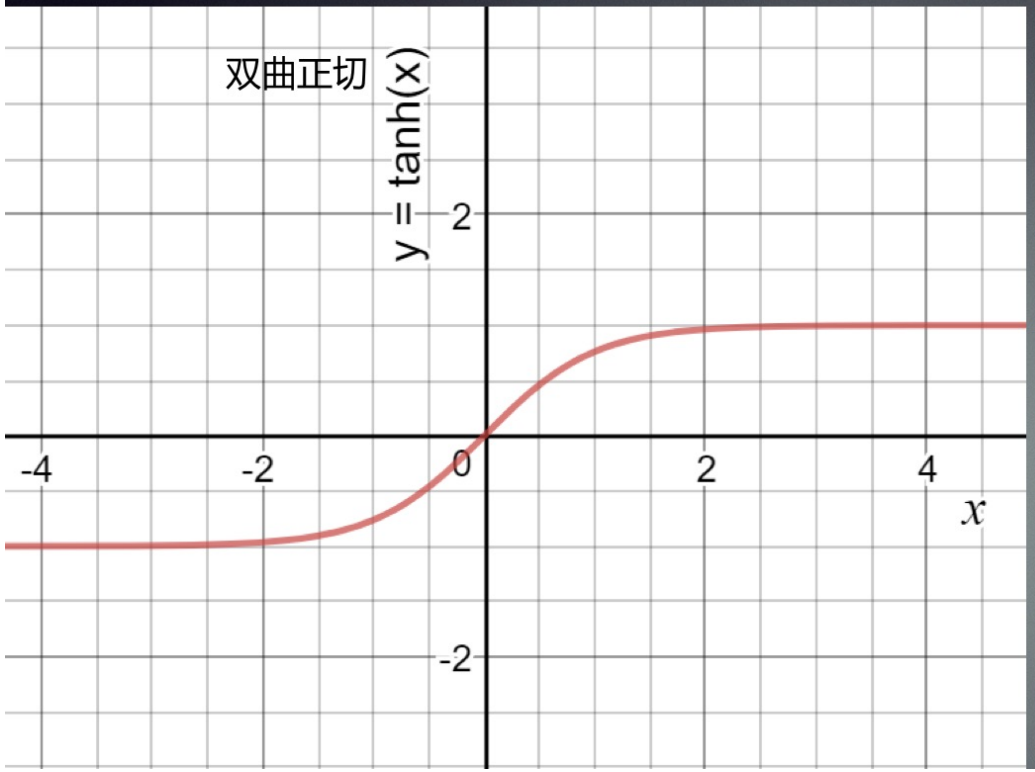

Nebraska
Lincoln
花萼
The sepals are located beneath the petals. They are often mistaken as petals. The sepals act as a protective layer around the unopened flower.


年龄
0.5
收入
0.8
性别
1
0.57
不买车的概率
0.43
买车的概率


0.5
0.8
1
0.1
0.2
0.3
0.4
0.1
0.2
0.3
0.5
0.7
0.2
0.1
0.3
0.4
1.23
A
0.84
0.3
0.76
A
0.64
0.6
1.09
A
0.80
0.1
1.0
A
0.76
0.5
0.6
0.4
0.7
0.2
0.1
0.2
0.7
0.9
2.08
1.81
Softmax
0.57
0.43
贪心科技 | 让每个人享受个性化教育服务

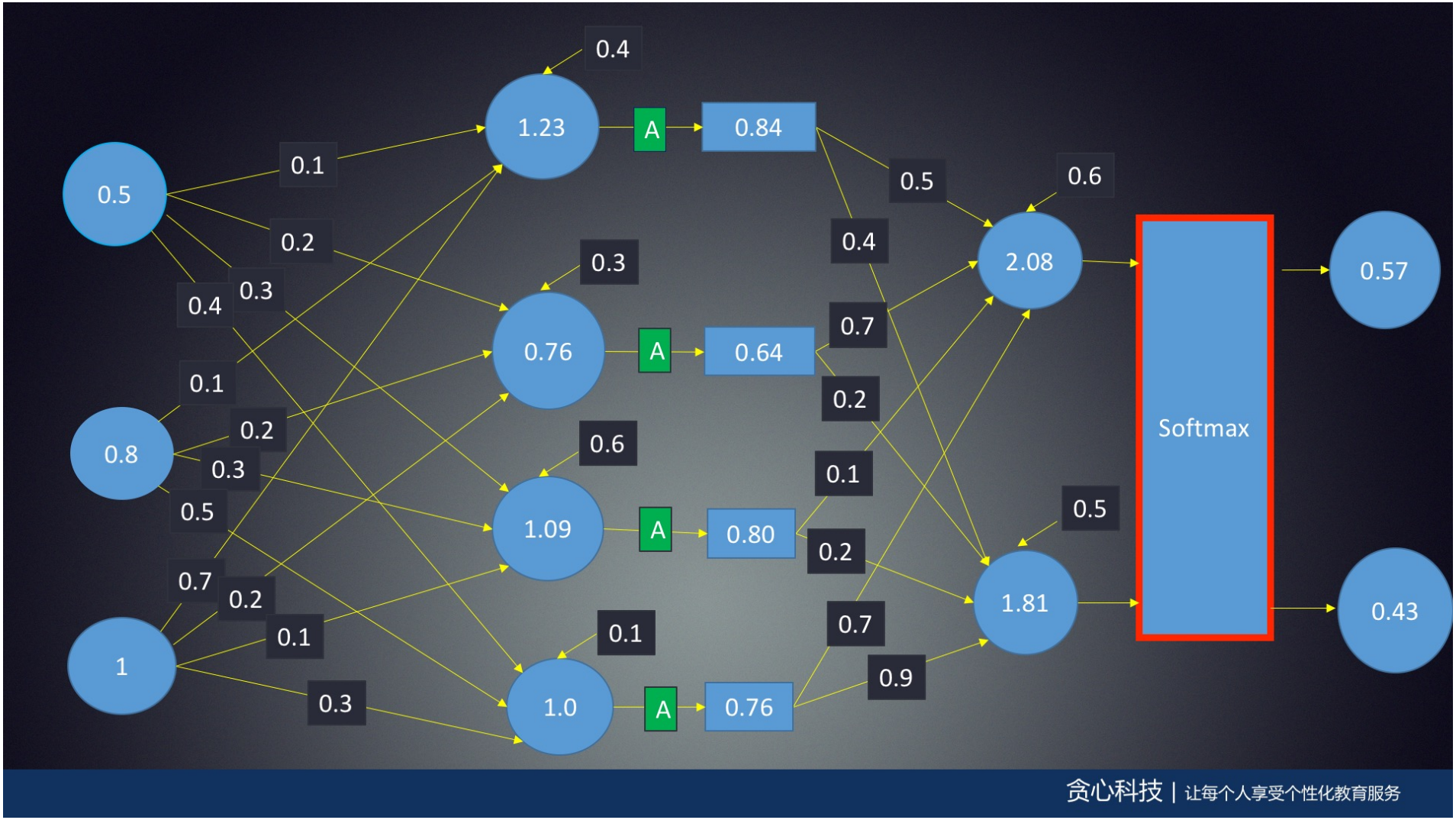

0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.4
2.08
0.57
0.2
0.3
0.3
0.4
0.7
0.76
A
0.64
0.1
0.2
0.2
Softmax
0.8
0.6
0.3
0.1
0.5
1.09
A
0.80
0.5
0.2
0.7
0.2
1.81
0.43
0.1
0.7
0.1
1
0.9
0.3
1.0
A
0.76

激活函数 Activation Function



$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \\ = \frac{e^{2x} - 1}{e^{2x} + 1} = \frac{1 - e^{-2x}}{1 + e^{-2x}}.$$

```
>>> import math
>>> math.tanh(1.9)
0.9562374581277391
>>>
>>> import numpy as np
>>> np.tanh([1.9, -1.9])
array([ 0.95623746, -0.95623746])
```


0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.4
2.08
0.57
0.2
0.3
0.3
0.4
0.7
0.76
A
0.64
0.1
0.2
0.2
Softmax
0.8
0.6
0.3
0.1
0.5
1.09
A
0.80
0.5
0.2
0.7
1.81
0.43
0.2
0.7
0.1
0.1
0.9
1
1.0
A
0.76
0.3

Softmax (Normalized exponential function, 归一化指数函数)

$$\sigma : \mathbb{R}^K \to \left\{ z \in \mathbb{R}^K | z_i > 0, \sum_{i=1}^{K} z_i = 1 \right\}$$

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

```
>>> import numpy as np
>>> def softmax(x):
...     return np.exp(x)/np.sum(np.exp(x))
...
>>> softmax([1, 2, 3])
array([ 0.09003057,  0.24472847,  0.66524096])
```

Softmax 数值问题 （overflow, 溢出）

```
>>> softmax([100, 200, 300])
array([  1.38389653e-87,   3.72007598e-44,   1.00000000e+00])

>>> softmax([1000, 2000, 3000])
__main__:2: RuntimeWarning: overflow encountered in exp
__main__:2: RuntimeWarning: invalid value encountered in true
_divide
array([ nan,  nan,  nan])
```

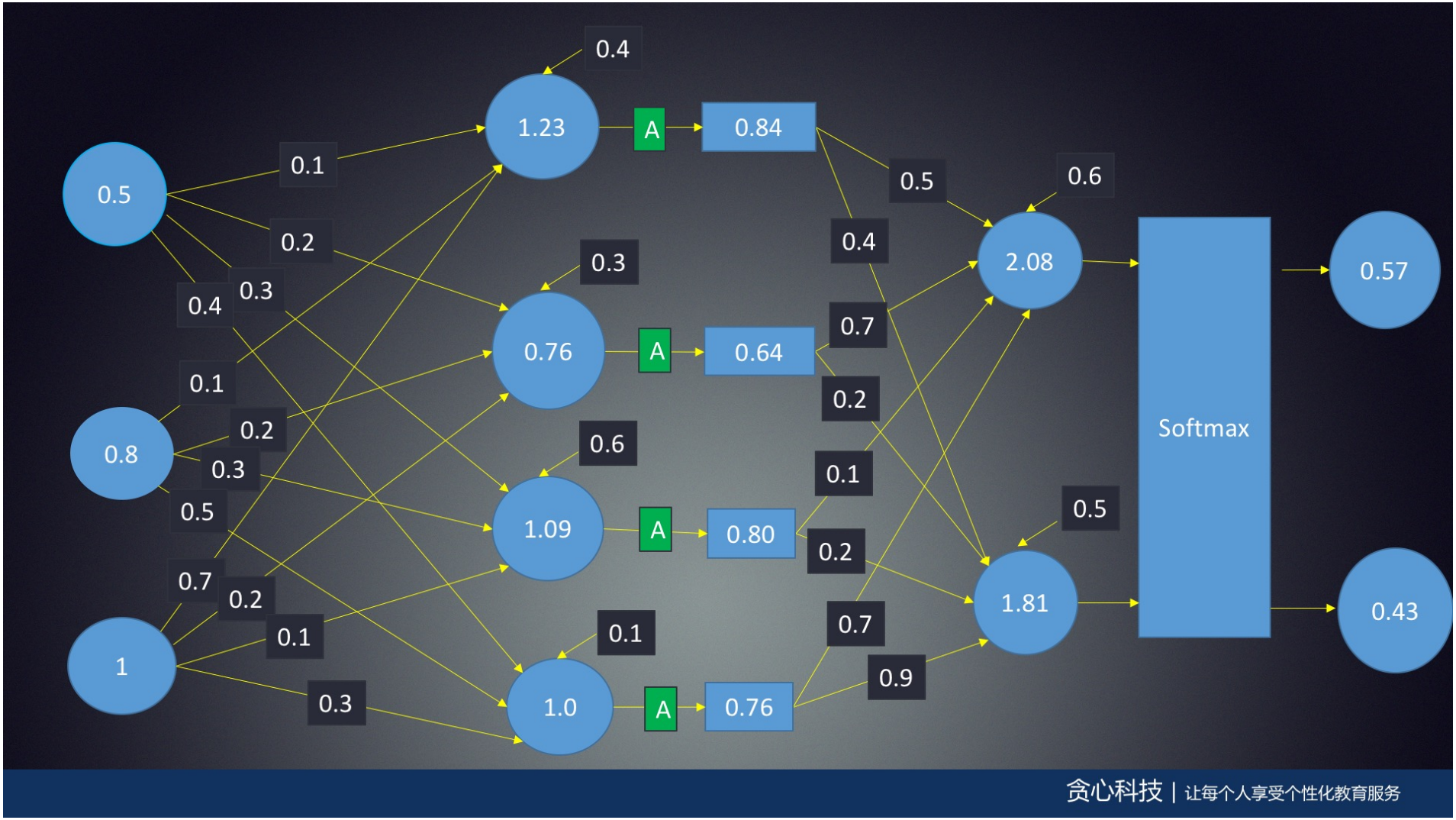
```
>>> def softmax(x):
...     m = np.max(x)
...     return np.exp(x-m)/np.sum(np.exp(x-m))
...
>>> softmax([1000, 2000, 3000])
array([ 0.,  0.,  1.])
>>> softmax([1, 2, 3])
array([ 0.09003057,  0.24472847,  0.66524096])
```

Softmax 数值问题 （overflow, 溢出）

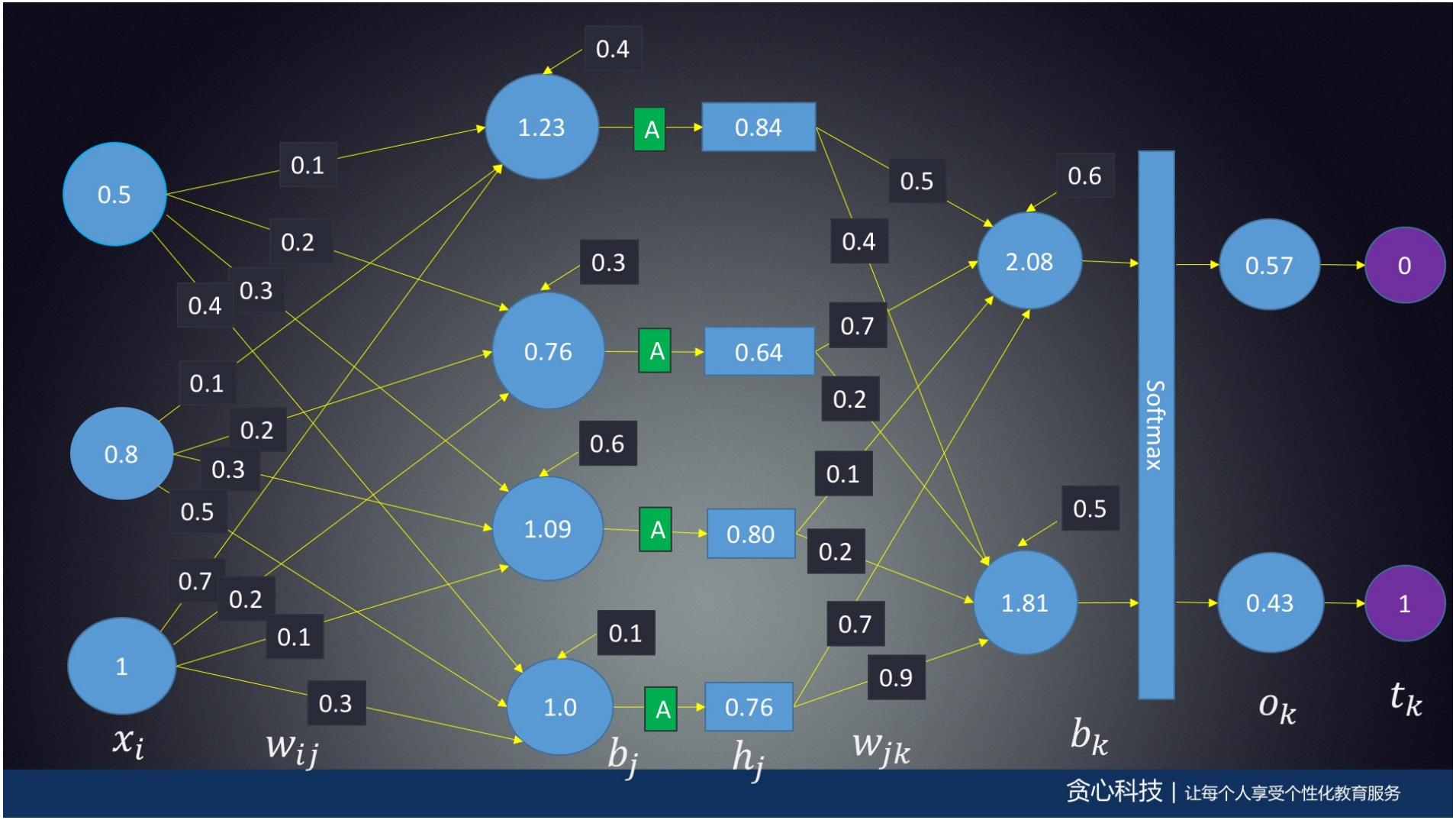
```
>>> softmax([100, 200, 300])
array([  1                                        .00000000e+00])

>>> softma
__main__:2                                        red in exp
__main__:2                                        untered in true
_divide
array([ nan,  nan,  nan])
```

$$\frac{e^{x_i}}{\sum_j e^{x_j}} = \frac{e^{-m}}{e^{-m}} \frac{e^{x_i}}{\sum_j e^{x_j}} = \frac{e^{x_i - m}}{\sum_j e^{x_j - m}}$$

```
>>> def softmax(x):
...     m = np.max(x)
...     return np.exp(x-m)/np.sum(np.exp(x-m))
...
>>> softmax([1000, 2000, 3000])
array([ 0.,  0.,  1.])
>>> softmax([1, 2, 3])
array([ 0.09003057,  0.24472847,  0.66524096])
```


0.5
0.8
1
0.1
0.2
0.3
0.4
0.1
0.2
0.3
0.5
0.7
0.2
0.1
0.3
0.4
1.23
A
0.84
0.3
0.76
A
0.64
0.6
1.09
A
0.80
0.1
1.0
A
0.76
0.5
0.6
0.4
0.7
0.2
0.1
0.2
0.7
0.9
2.08
1.81
Softmax
0.57
0.43
贪心科技 | 让每个人享受个性化教育服务

误差反向传播（BP算法）



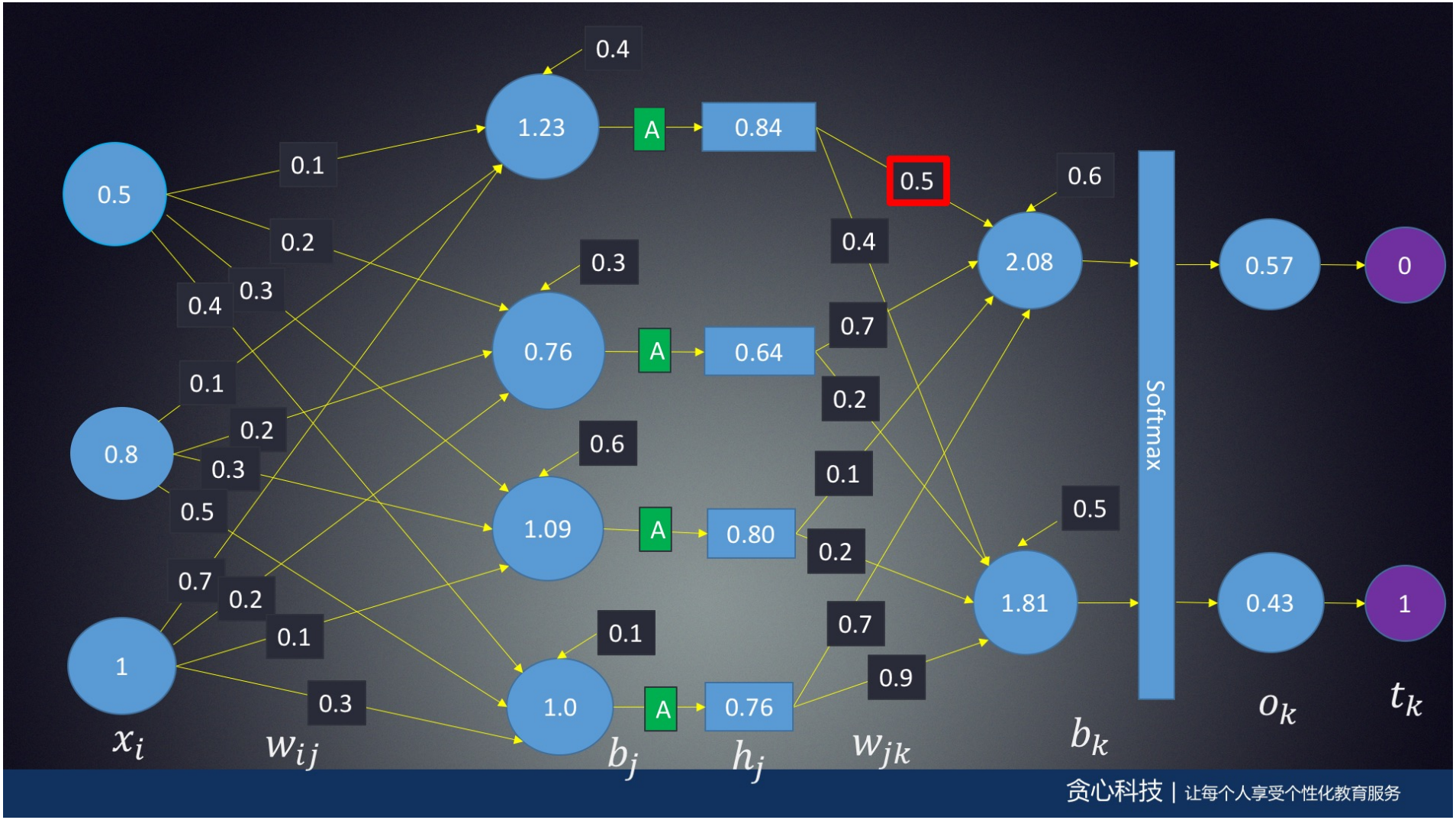
$$L = \frac{1}{2}(t_k - o_k)^2$$

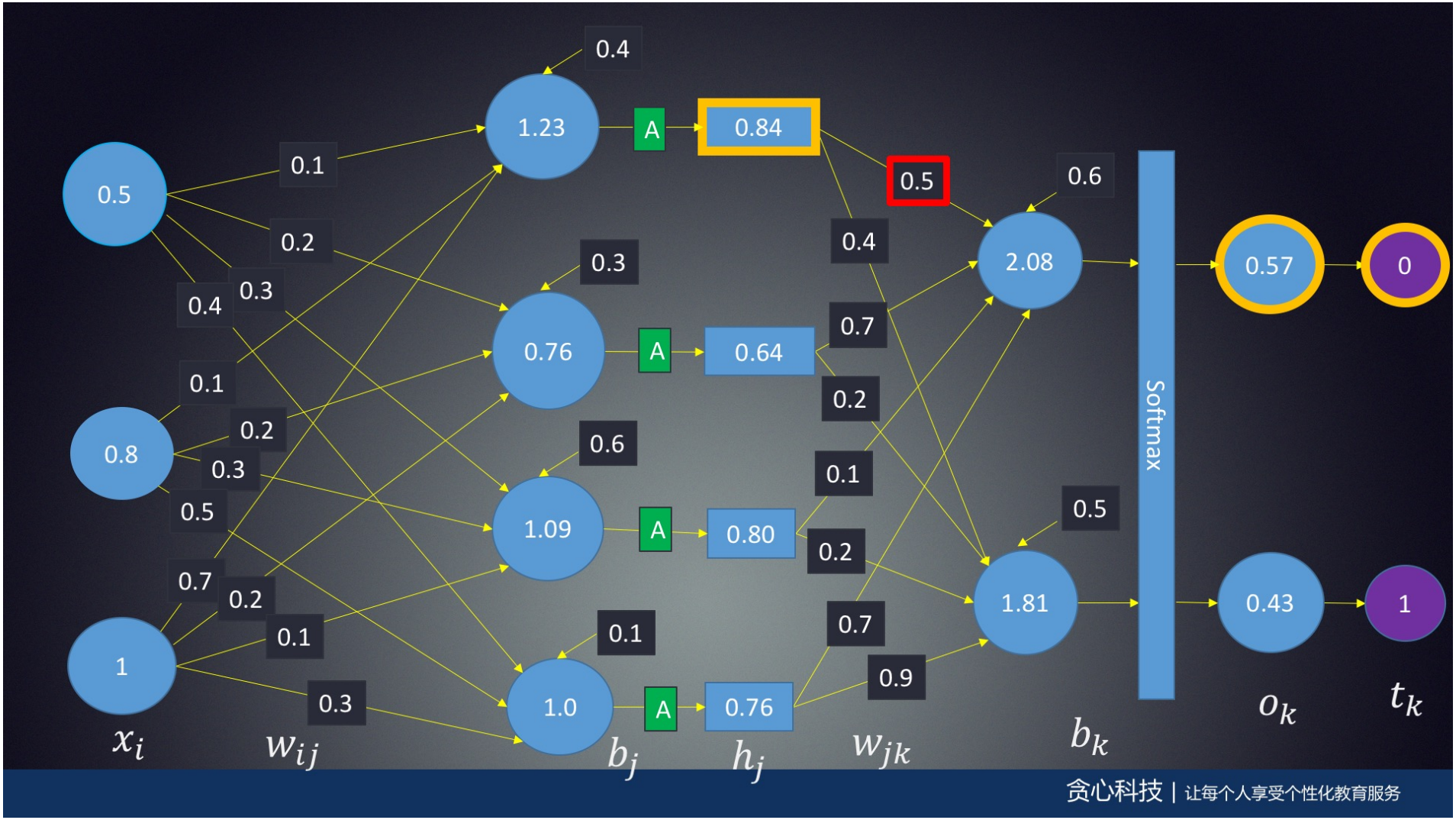
$$\frac{\partial L}{\partial w_{jk}} = h_j \underbrace{(o_k - t_k)o_k(1 - o_k)}_{\delta_k}$$

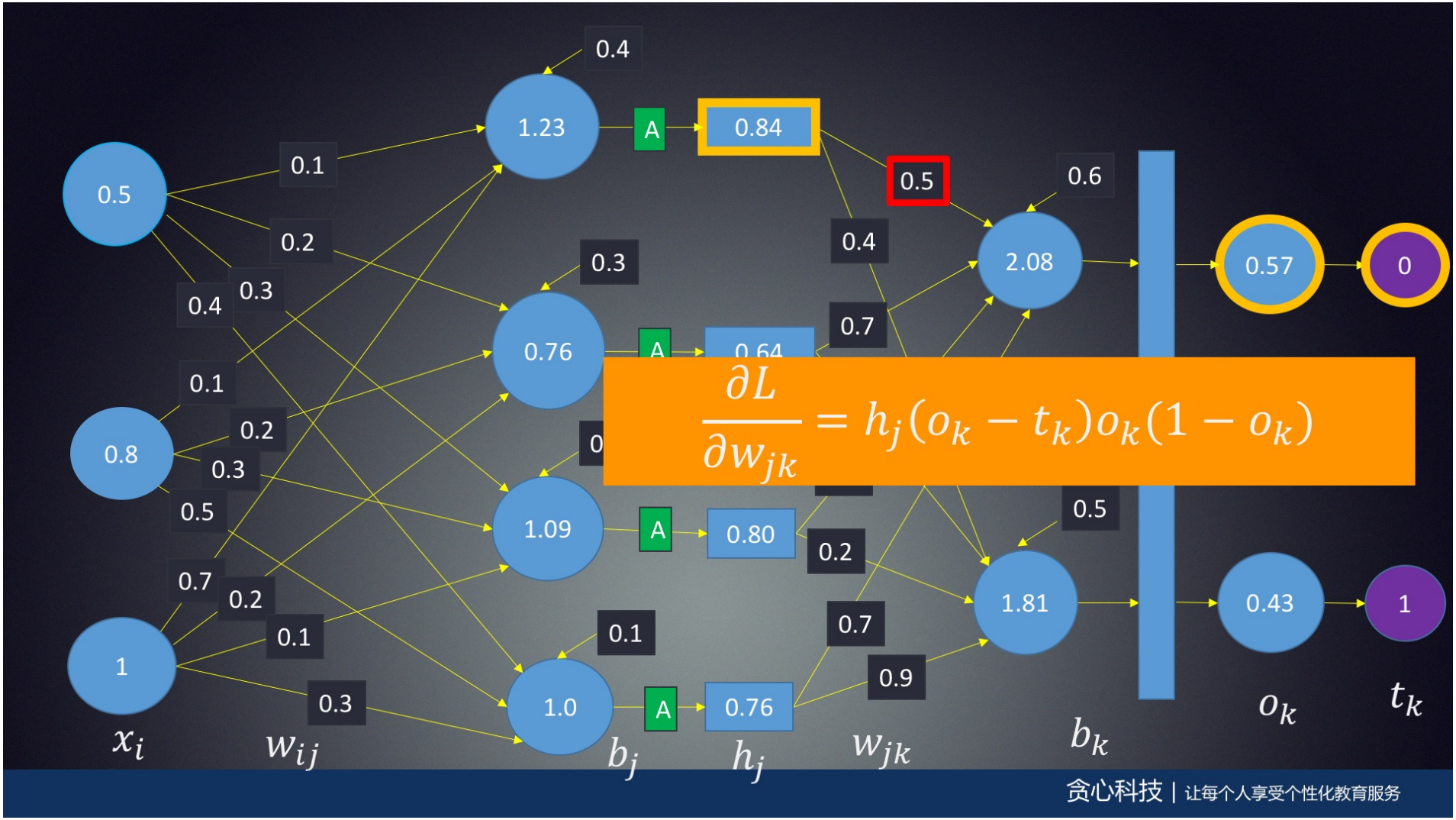
$$\frac{\partial L}{\partial b_k} = (o_k - t_k)o_k(1 - o_k)$$

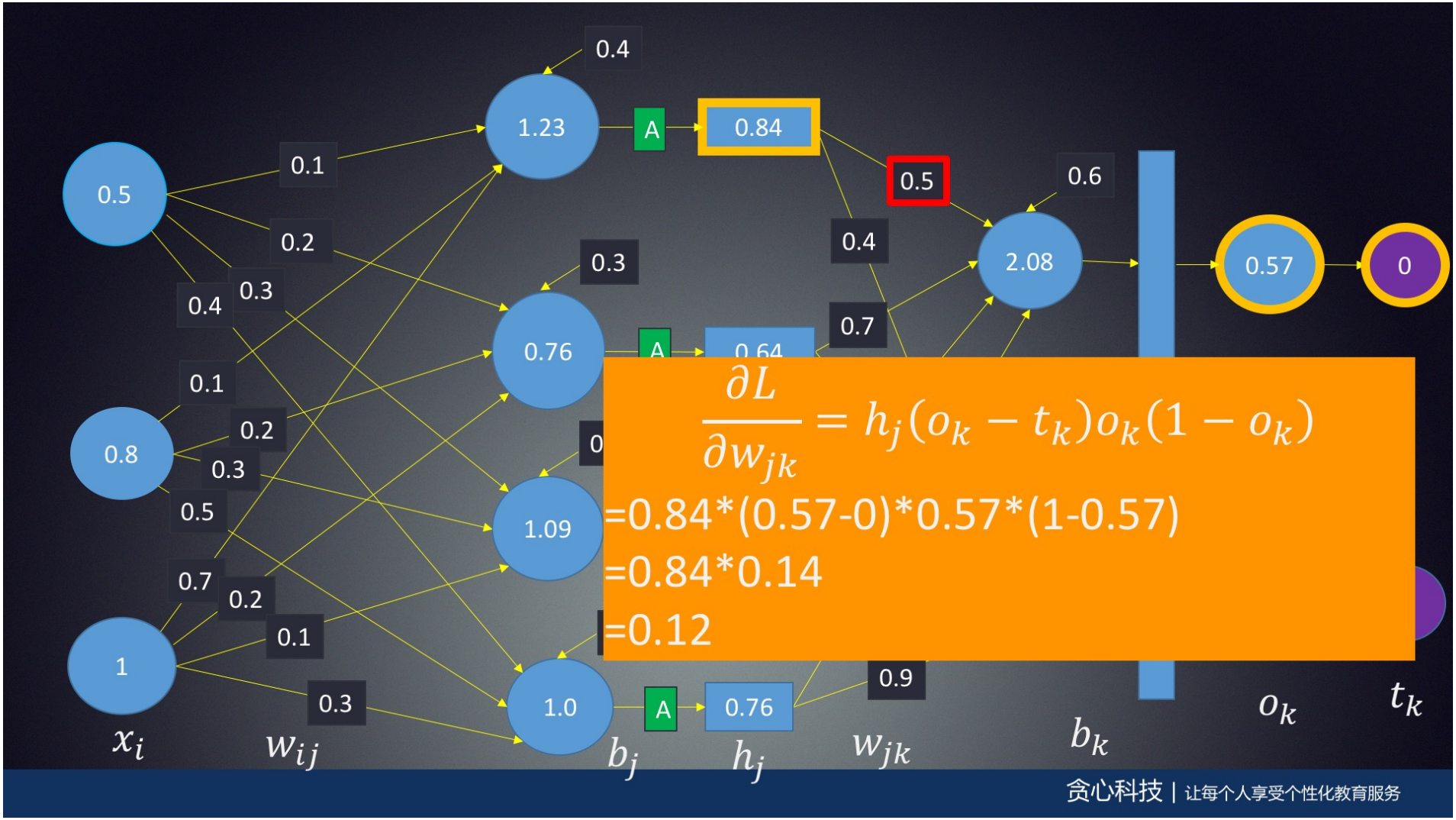
$$\frac{\partial L}{\partial w_{ij}} = x_i \sum_k \delta_k w_{jk} \left(1 + h_j\right)\left(1 - h_j\right)$$

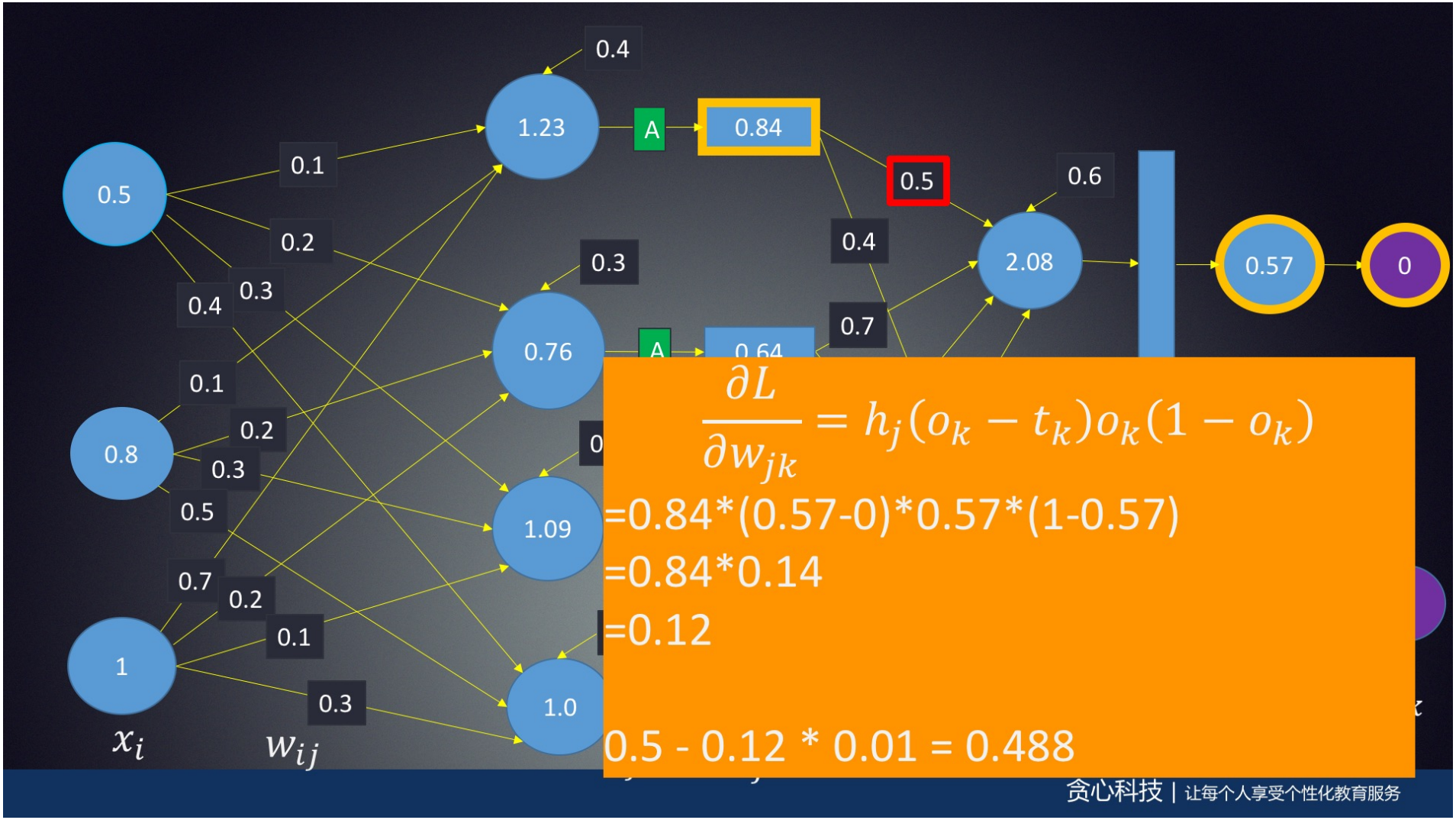
$$\frac{\partial L}{\partial b_j} = \sum_k \delta_k w_{jk} \left(1 + h_j\right)\left(1 - h_j\right)$$

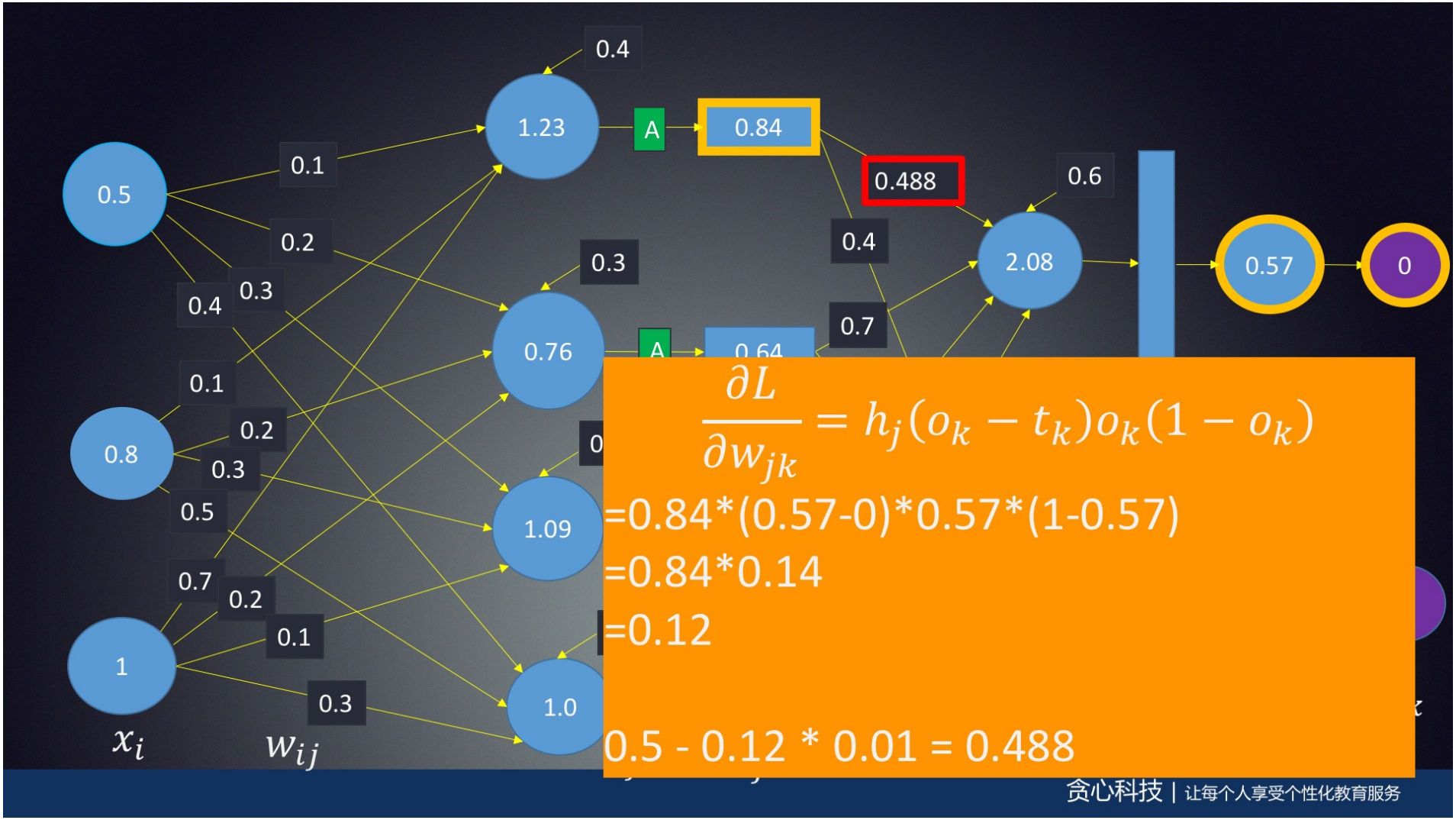

0.5
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.1
0.2
0.3
0.4
0.1
0.2
0.3
0.6
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.9
1.23
0.76
1.09
1.0
A
0.84
0.64
0.80
0.76
2.08
1.81
Softmax
0.57
0.43
0
1
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
贪心科技 | 让每个人享受个性化教育服务

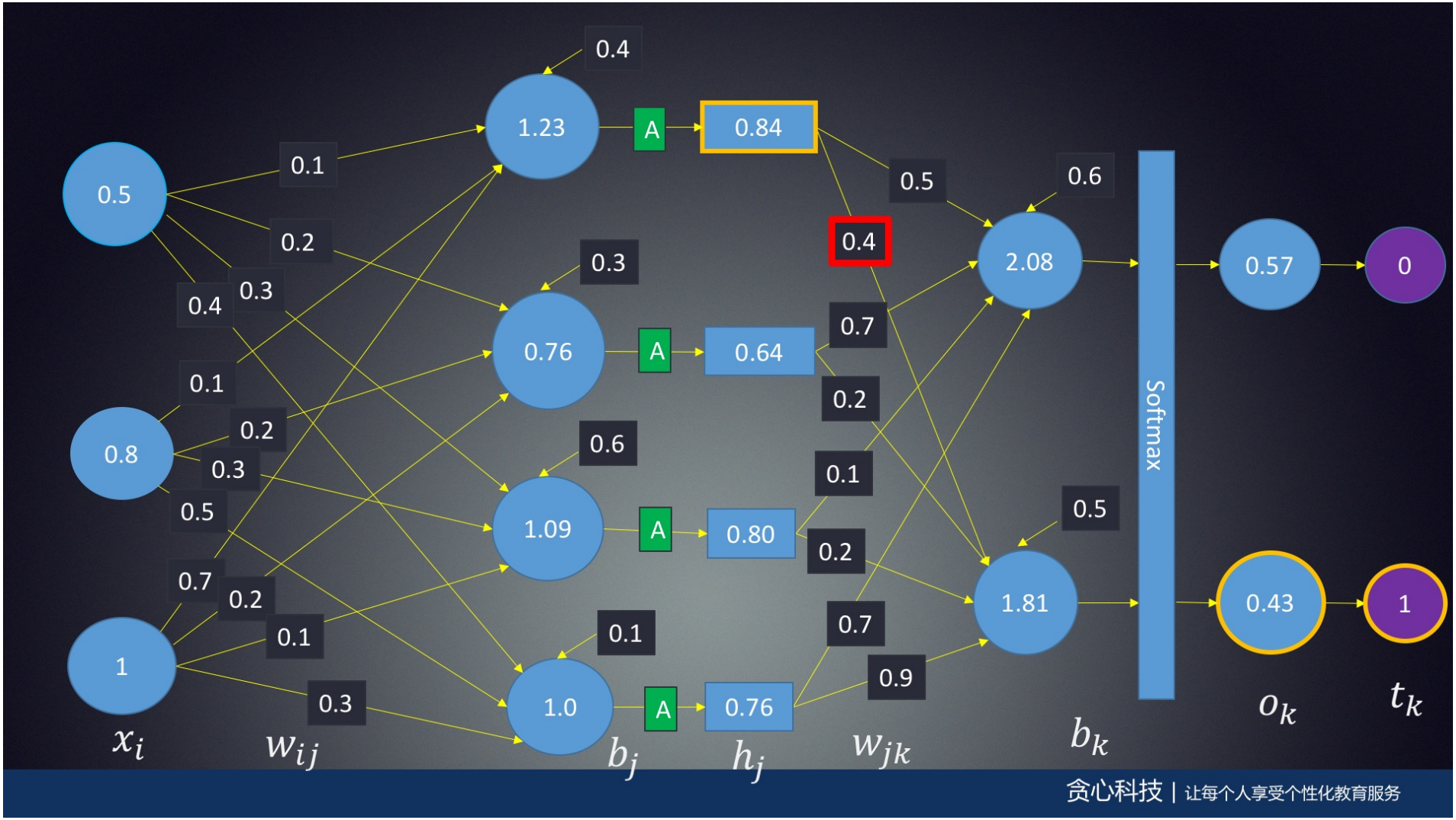

0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.2
0.4
0.3
2.08
0.57
0
0.3
0.4
0.7
0.76
A
0.64
0.1
0.2
0.8
0.2
0.6
0.3
0.1
0.5
1.09
A
0.80
0.2
0.5
0.7
0.2
1.81
0.43
1
0.1
0.1
0.7
1
0.9
0.3
1.0
A
0.76
Softmax
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
贪心科技 | 让每个人享受个性化教育服务

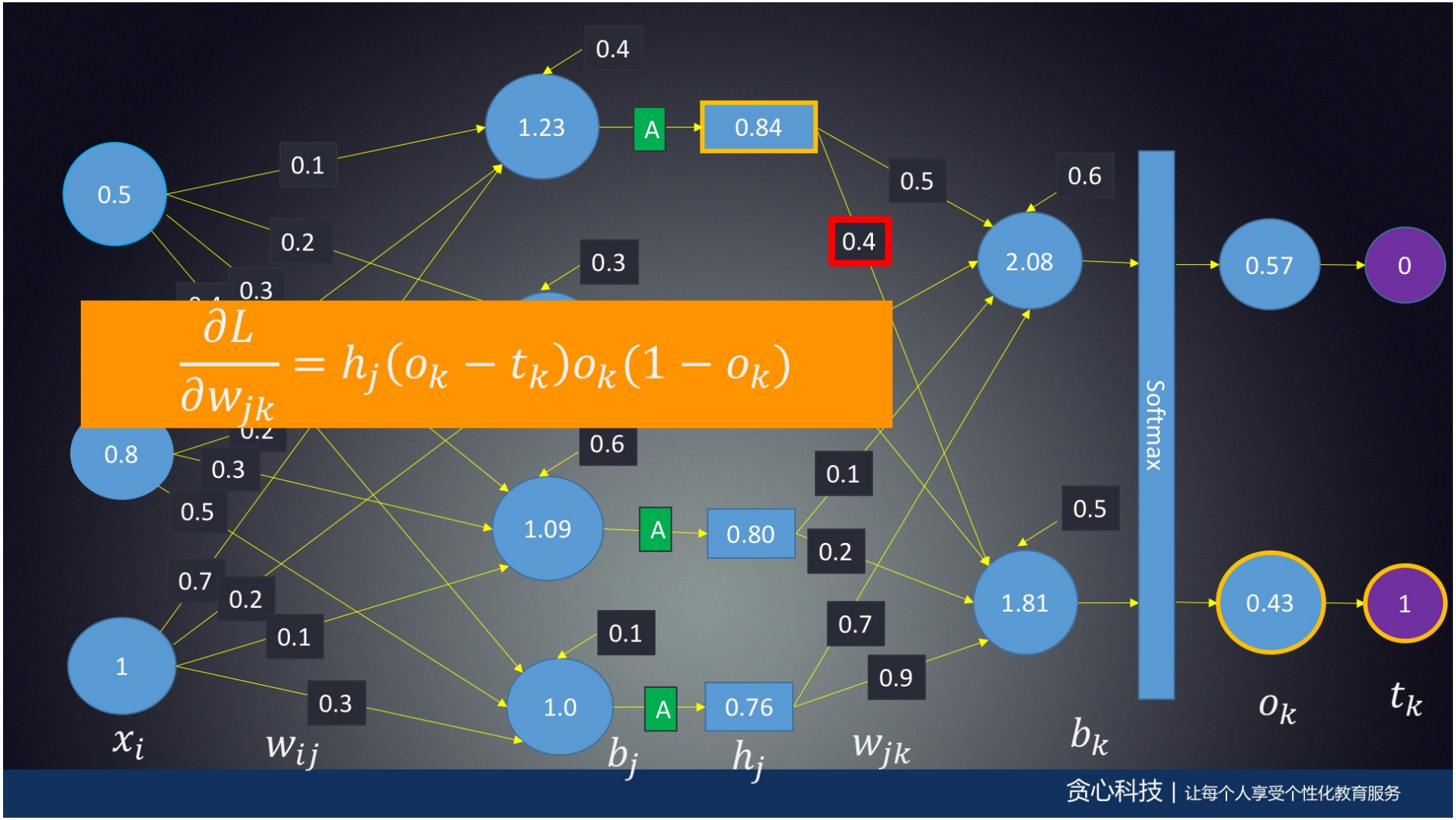

0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.2
0.4
0.3
0.4
2.08
0.57
0
0.3
0.7
0.76
A
0.64
0.1
0.2
0.8
0.3
0.5
1.09
A
0.80
0.5
0.2
0.7
0.2
1.81
0.43
1
0.1
0.1
0.7
1
0.3
0.9
1.0
A
0.76
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$

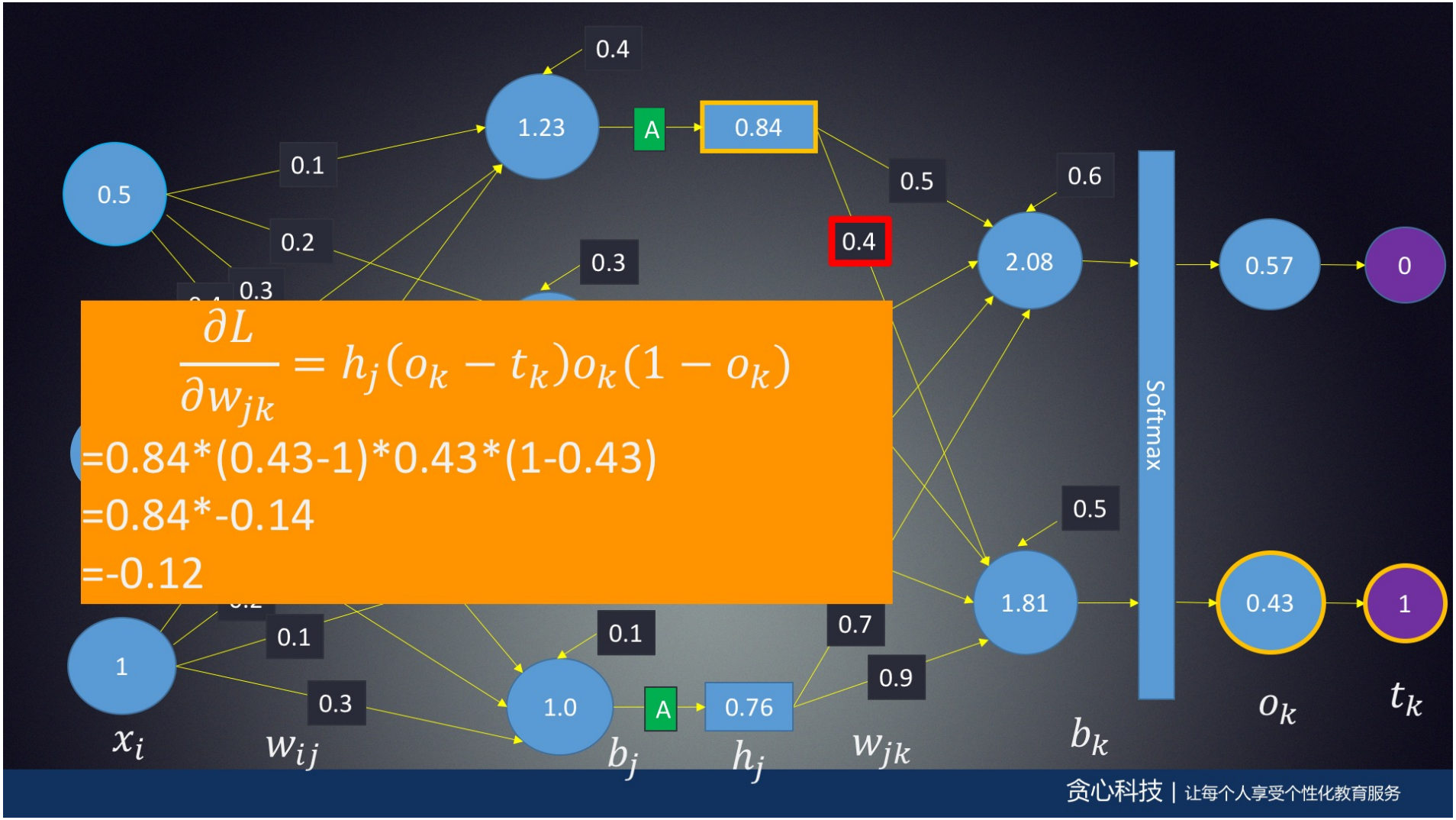

0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.2
0.4
0.3
0.4
2.08
0.57
0
0.3
0.7
0.76
A
0.64
0.1
0.2
0.8
0.3
0.5
1.09
0.7
0.2
0.1
1
0.3
1.0
A
0.76
0.9
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
$$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$$
=0.84*(0.57-0)*0.57*(1-0.57)
=0.84*0.14
=0.12

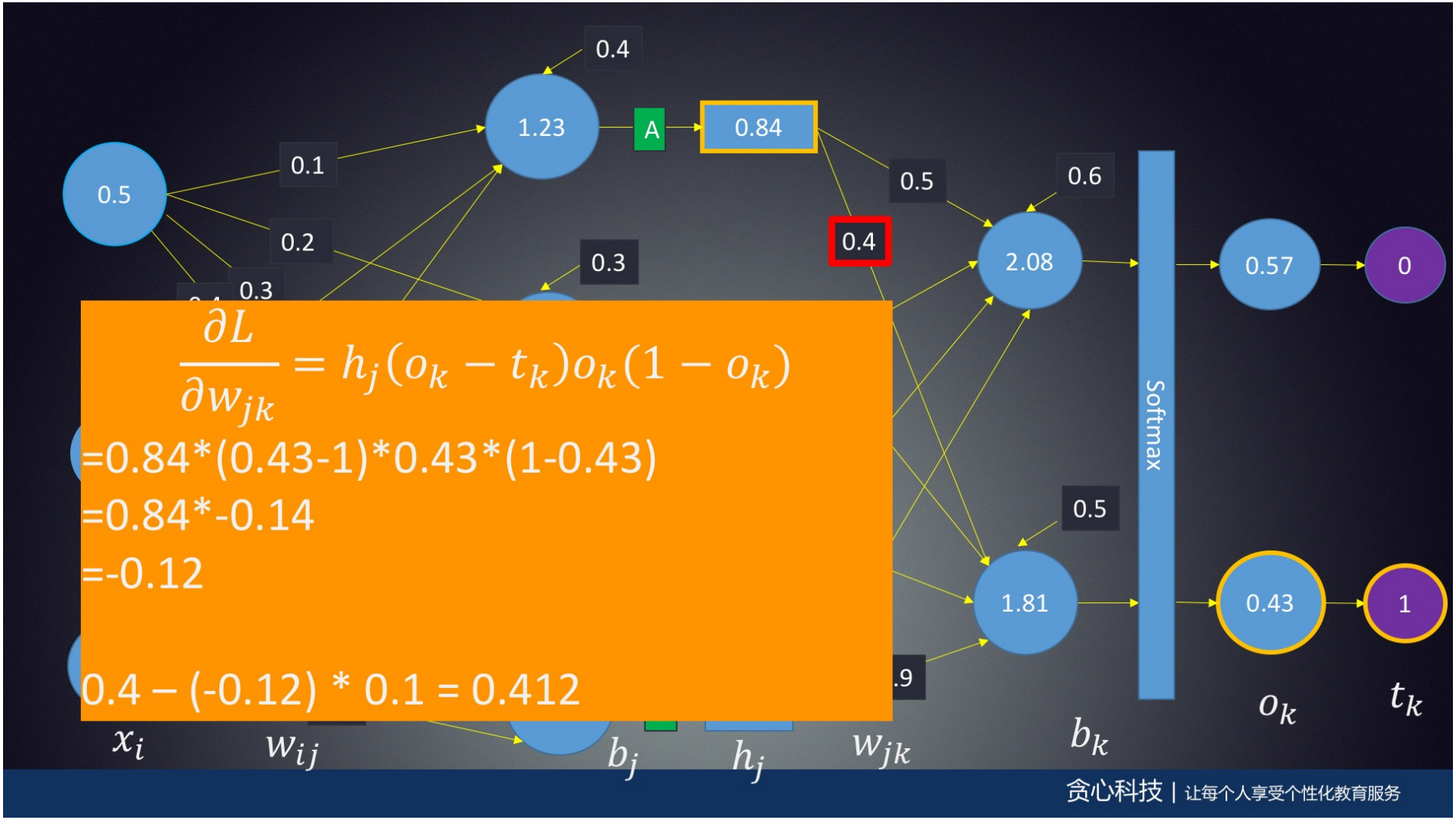

0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.2
0.4
0.3
0.4
2.08
0.57
0
0.3
0.7
0.76
A
0.64
0.1
0.2
0.8
0.3
0.5
1.09
0.7
0.2
0.1
1
0.3
1.0
x_i
w_{ij}
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$
=0.84*(0.57-0)*0.57*(1-0.57)
=0.84*0.14
=0.12
0.5 - 0.12 * 0.01 = 0.488

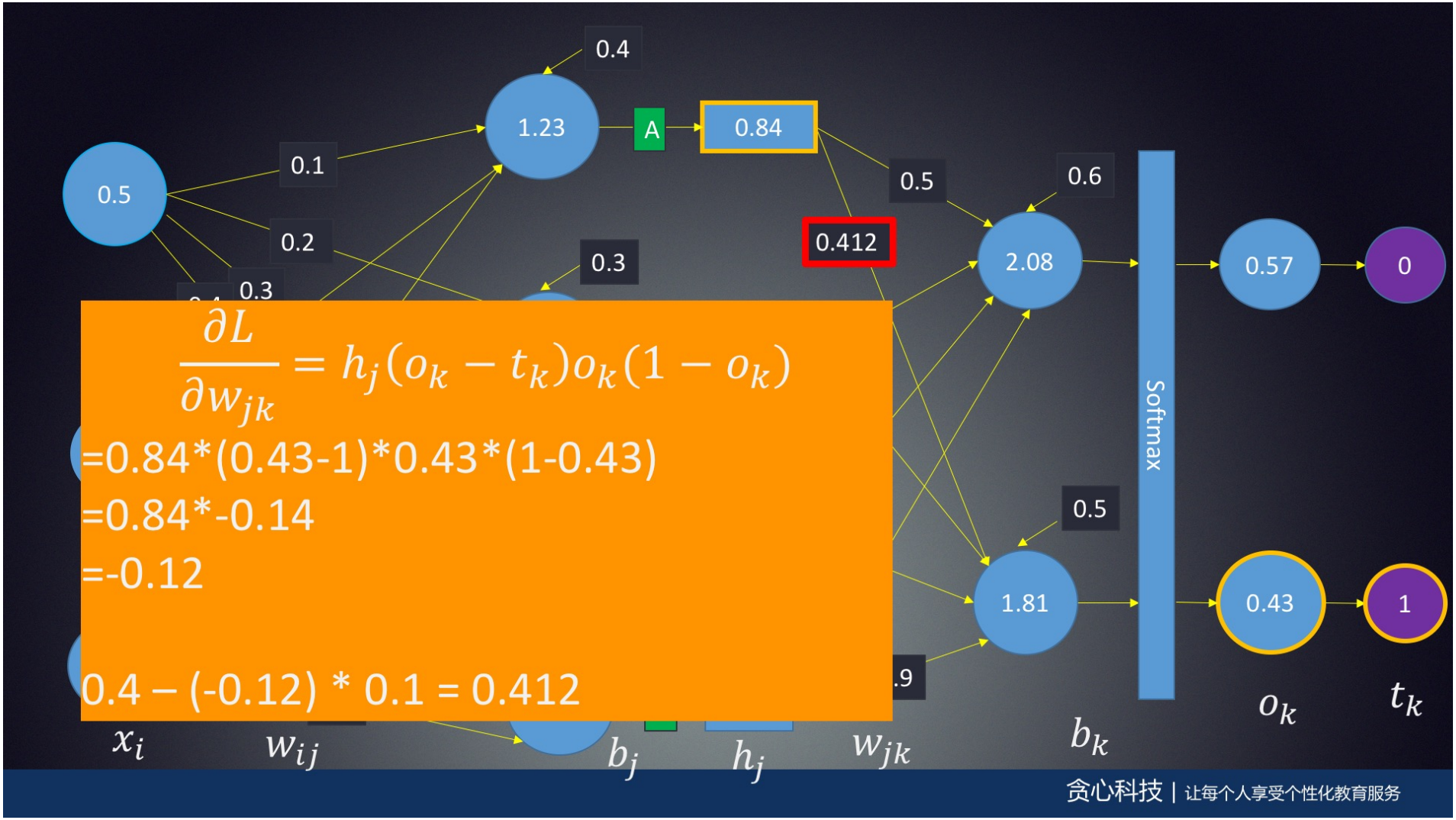

0.4
1.23
A
0.84
0.488
0.6
0.5
0.1
0.2
0.3
0.4
0.4
2.08
0.57
0
0.3
0.76
0.7
0.1
0.8
0.2
0.3
0.5
1.09
0.7
0.2
0.1
1
0.3
1.0
x_i
w_{ij}
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$
=0.84*(0.57-0)*0.57*(1-0.57)
=0.84*0.14
=0.12
0.5 - 0.12 * 0.01 = 0.488

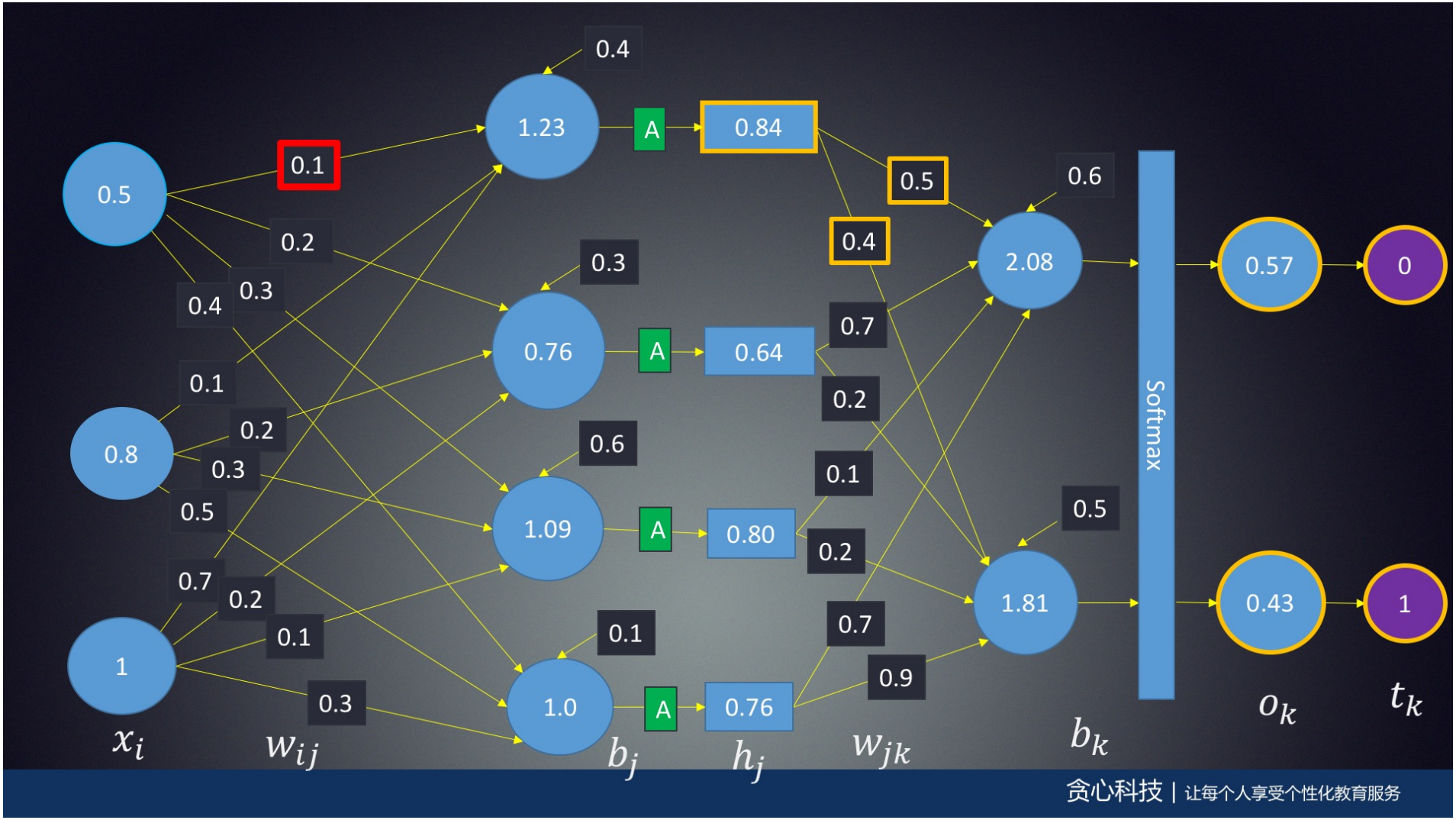

0.4
1.23
A
0.84
0.5
0.6
0.1
0.5
0.4
2.08
0.57
0
0.2
0.3
0.3
0.4
0.7
0.76
A
0.64
0.1
0.2
0.2
0.8
0.6
0.3
0.1
0.5
0.5
1.09
A
0.80
0.2
0.7
0.2
1.81
0.43
1
0.1
0.7
0.1
0.9
1
0.3
1.0
A
0.76
Softmax
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k

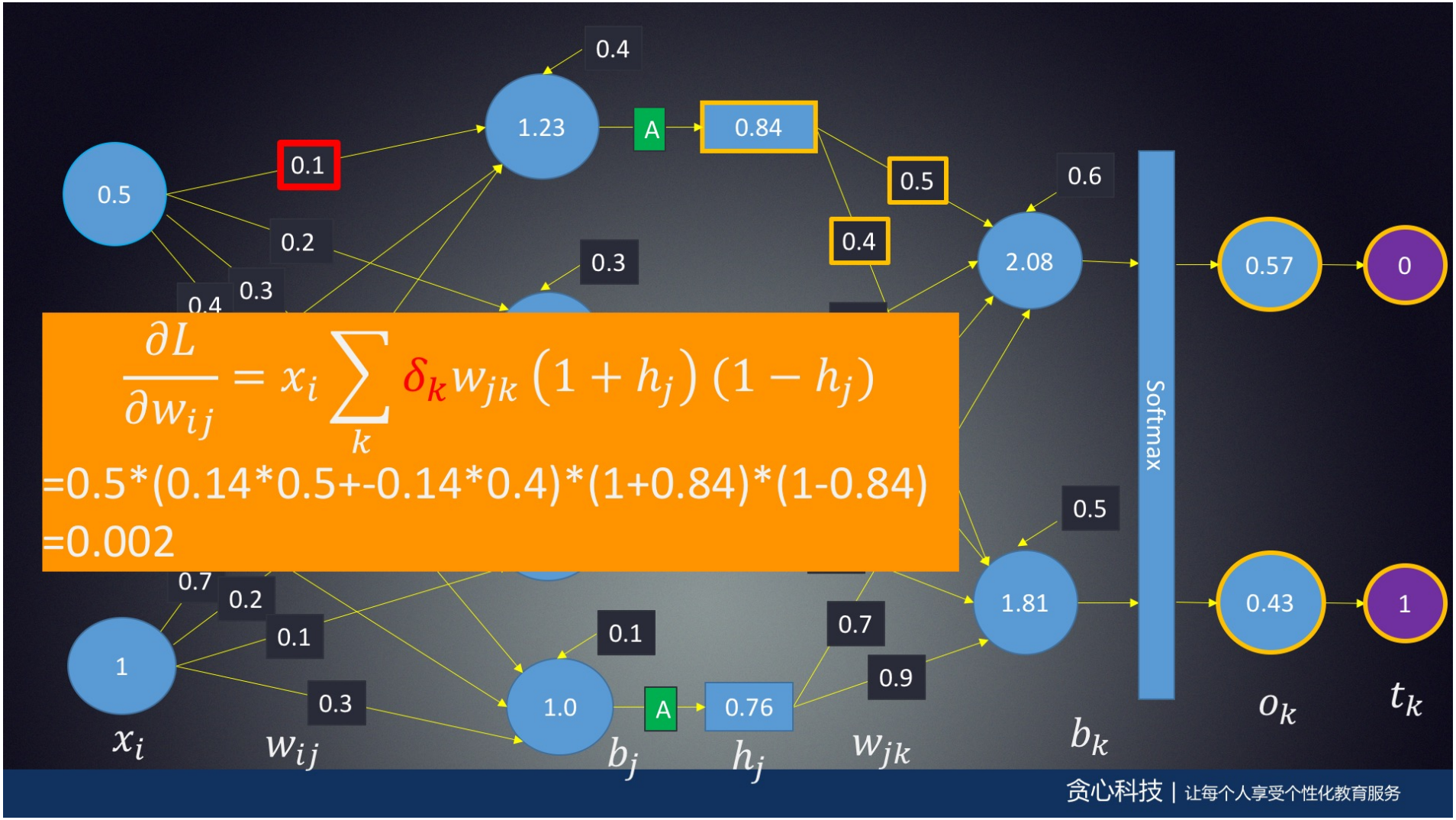

0.4
1.23
A
0.84
0.5
0.1
0.2
0.3
0.4
0.6
2.08
Softmax
0.57
0
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$
0.8
0.3
0.5
0.6
1.09
A
0.80
0.1
0.2
0.5
0.7
0.2
0.1
1.81
0.43
1
0.1
0.7
0.9
1
0.3
1.0
A
0.76
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
贪心科技 | 让每个人享受个性化教育服务

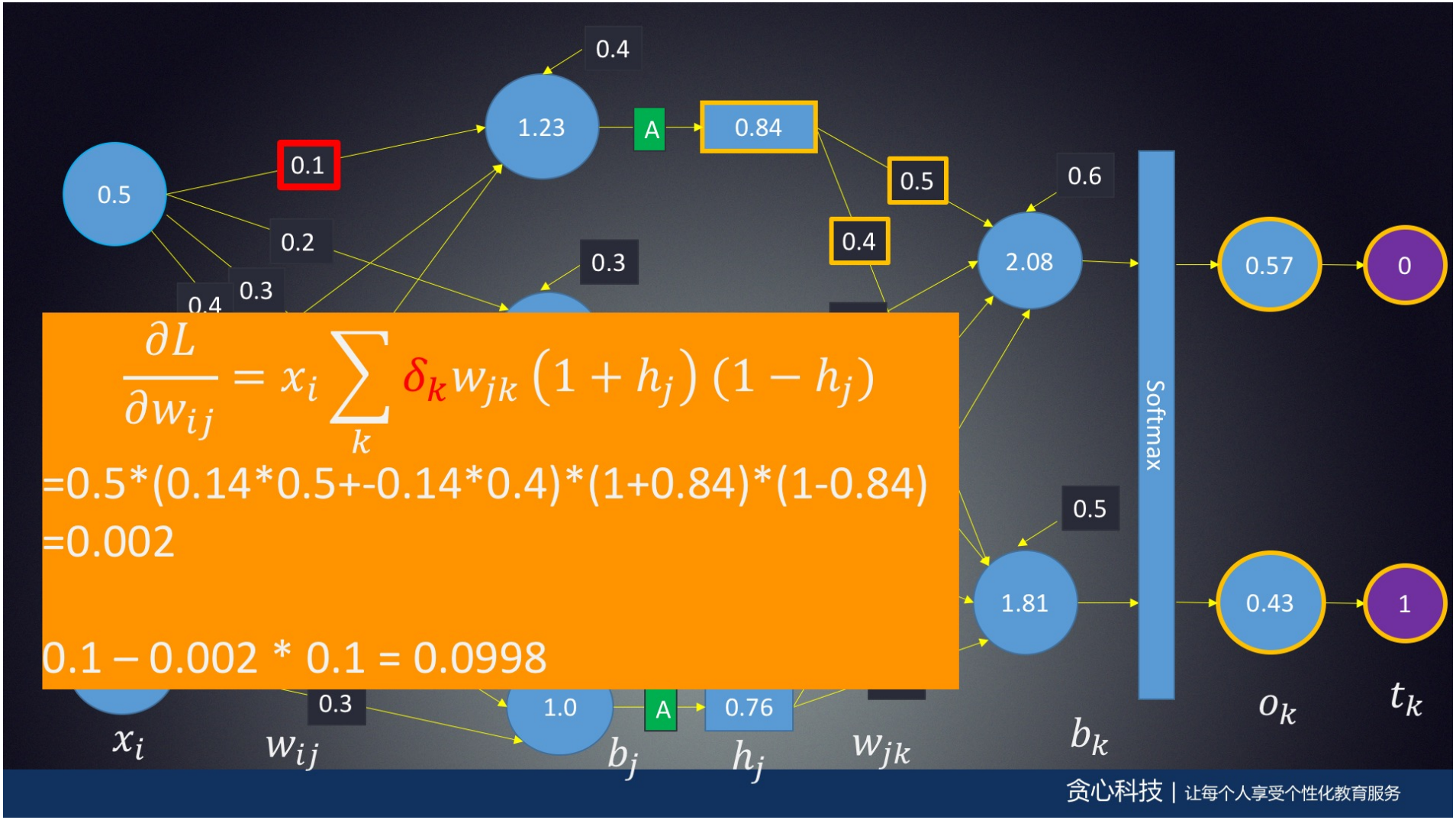

0.4
1.23
A
0.84
0.5
0.1
0.2
0.3
0.4
0.6
2.08
Softmax
0.57
0
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$
=0.84*(0.43-1)*0.43*(1-0.43)
=0.84*-0.14
=-0.12
0.5
1.81
0.43
1
1
0.1
0.1
0.7
0.9
0.3
1.0
A
0.76
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k


0.4
1.23
A
0.84
0.5
0.1
0.2
0.3
0.4
0.6
2.08
Softmax
0.57
0
0.5
1.81
0.43
1
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$
=0.84*(0.43-1)*0.43*(1-0.43)
=0.84*-0.14
=-0.12
0.4 – (-0.12) * 0.1 = 0.412
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k


0.4
1.23
A
0.84
0.5
0.1
0.2
0.3
0.412
0.6
2.08
Softmax
0.57
0
0.5
1.81
0.43
1
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
$\frac{\partial L}{\partial w_{jk}} = h_j(o_k - t_k)o_k(1 - o_k)$
=0.84*(0.43-1)*0.43*(1-0.43)
=0.84*-0.14
=-0.12
0.4 – (-0.12) * 0.1 = 0.412
贪心科技 | 让每个人享受个性化教育服务


0.5
0.8
1
0.1
0.2
0.3
0.4
0.5
0.7
0.2
0.1
0.3
0.4
0.1
0.2
0.3
0.5
0.6
0.1
0.2
0.3
0.7
0.2
0.1
0.4
1.23
0.76
1.09
1.0
A
A
A
A
0.84
0.64
0.80
0.76
0.5
0.4
0.7
0.2
0.1
0.2
0.7
0.9
0.6
0.5
2.08
1.81
Softmax
0.57
0.43
0
1
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k
贪心科技 | 让每个人享受个性化教育服务


0.4
1.23
A
0.84
0.5
0.1
0.5
0.6
0.2
0.4
2.08
0.57
0
0.3
0.4
0.3
Softmax
$\frac{\partial L}{\partial w_{ij}} = x_i \sum_k \delta_k w_{jk} (1 + h_j)(1 - h_j)$
=0.5*(0.14*0.5+-0.14*0.4)*(1+0.84)*(1-0.84)
=0.002
0.5
0.7
0.2
1.81
0.43
1
0.1
0.1
0.7
1
0.9
0.3
1.0
A
0.76
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k

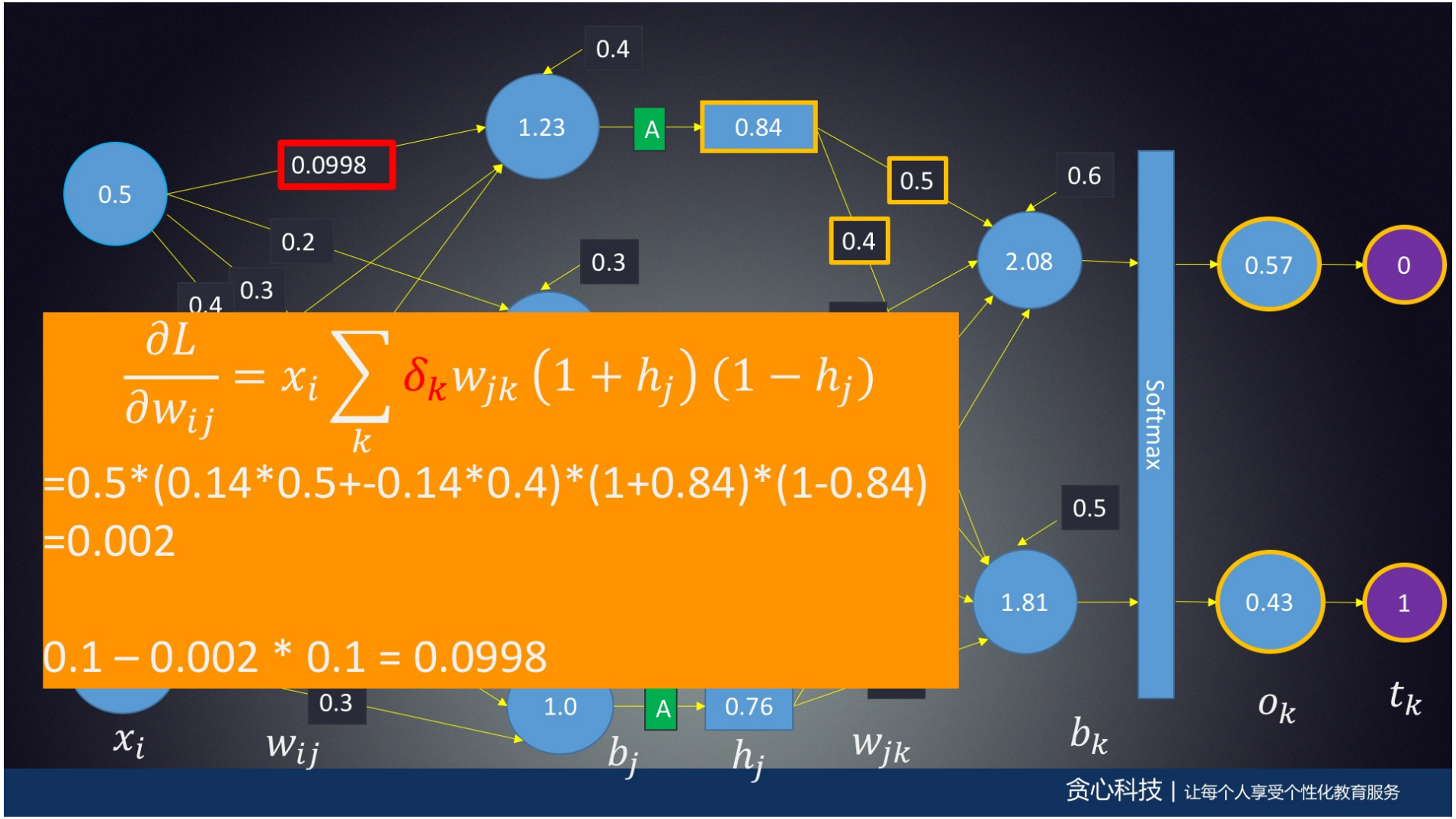


$$\frac{\partial L}{\partial w_{ij}} = x_i \sum_k \delta_k w_{jk} \left(1 + h_j\right) \left(1 - h_j\right)$$

=0.5*(0.14*0.5+-0.14*0.4)*(1+0.84)*(1-0.84)

=0.002

0.1 – 0.002 * 0.1 = 0.0998


0.4
1.23
A
0.84
0.5
0.0998
0.5
0.2
0.3
0.4
0.4
0.3
0.6
2.08
Softmax
0.57
0
$\frac{\partial L}{\partial w_{ij}} = x_i \sum_k \delta_k w_{jk} (1 + h_j)(1 - h_j)$
=0.5*(0.14*0.5+-0.14*0.4)*(1+0.84)*(1-0.84)
=0.002
0.1 – 0.002 * 0.1 = 0.0998
0.5
1.81
0.43
1
0.3
1.0
A
0.76
x_i
w_{ij}
b_j
h_j
w_{jk}
b_k
o_k
t_k