1.

$B(\frac{9}{18}) = 1$

s: skips

r: reads

3.

Length

long → e1, e3, e4, e6, e9, e10, e12 (s=7, r=0)

short → e2 e5 e7 e8 e11 e13 e14 e15 e16 e17 e18 (s=2, r=9)

$B(0) = 0$

$B(\frac{9}{11}) = -\frac{9}{11}\log_2\frac{9}{11} - \frac{2}{11}\log_2\frac{2}{11}$

① choose Author:

$Gain(Length) = 1 - \frac{11}{18}B(\frac{9}{11}) \approx 0.5820$

Author

known → e5 e13 e14 e16 e17 e15 (s=0, r=6)

unknown → e2 e7 e8 e18 e11 (s=2, r=3)

$Gain(Author) = B(\frac{9}{11}) - \frac{6}{11}B(\frac{6}{6}) - \frac{5}{11}B(\frac{3}{5})$

≈ 0.24270

② Choose Thread

Thread

new → e2 e5 e8 e14 e15 e17 e18 (s=0, r=7)

follow Up → e7 e11 e16 e13 (s=2, r=2)

$Gain(Thread) = B(\frac{9}{11}) - \frac{7}{11}B(\frac{7}{7}) - \frac{4}{11}B(\frac{1}{2})$

≈ 0.33040

③ Choose Where Read

Where Read

work → e2 e7 e8 e14 e16 e18 (s=1, r=5)

home → e5 e11 e13 e15 e17 (s=1, r=4)

$Gain(Where\ Read)$

$= B(\frac{9}{11}) - \frac{6}{11}B(\frac{5}{6}) - \frac{5}{11}B(\frac{4}{5})$

$= 0.00133$

故第二层选择 Thread

2.

解：(a)

A	B	C	D	E
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	1
0	1	1	0	1
0	1	1	1	0
1	0	0	0	0
1	0	0	1	1
1	0	1	0	1
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	0

(b) $P(E|A,B,C,D) = \frac{P(A,B,C,D|E)P(E)}{P(A,B,C,D)} = \alpha P(E)P(A|E)P(B|E)P(C|E)P(D|E)$,

其中 $\alpha = \frac{1}{P(A,B,C,D)}$ 为正常数．则：

$$\begin{aligned} P(e|a,b,c,d) &= \alpha P(e)P(a|e)P(b|e)P(c|e)P(d|e) \\ &= \alpha \times \frac{4}{15} \times \frac{2}{4} \times \frac{2}{4} \times \frac{2}{4} \times \frac{2}{4} \\ &= \frac{1}{60}\alpha \approx 0.0167\alpha \end{aligned}$$

$$\begin{aligned} P(\neg e|a,b,c,d) &= \alpha P(\neg e)P(a|\neg e)P(b|\neg e)P(c|\neg e)P(d|\neg e) \\ &= \alpha \times \frac{11}{15} \times \frac{5}{11} \times \frac{5}{11} \times \frac{5}{11} \times \frac{5}{11} \\ &= \frac{125}{3993}\alpha \approx 0.0313\alpha \end{aligned}$$

$\because P(\neg e|a,b,c,d) > P(e|a,b,c,d)$

$\therefore$ 预测此时 $E=0$.

感谢窦正同学的答案

3. 由 ML24-4 第 35 页和 36 页的推导，有

$$\frac{\partial Loss_{k=k_0}}{\partial w_{j=j_0,k=k_0}} = -2(y_{k_0} - a_{k_0})g'(\sum_j w_{j,k_0}a_j)a_{j_0}$$

$$\frac{\partial Loss_{k=k_0}}{\partial w_{i=i_0,j=j_0}} = -2(y_{k_0} - a_{k_0})g'(\sum_j w_{j,k_0}a_j)w_{j_0,k_0}g'(\sum_i w_{i,j_0}a_i)a_{i_0}$$

其中，$in_j = \Sigma_j w_{i,j}a_i = \Sigma_j w_{i,j}x_i$，$in_k = \sum_k w_{j,k}a_j = \sum_k w_{j,k}g(in_j)$。

(a) 要计算 $\frac{\partial Loss_{o_1}}{\partial w_1}$，即是要计算 $\frac{\partial Loss_{k=1}}{\partial w_{i=1,j=1}}$。从课件数据计算式中各项，可得

- $y_1 = 0.01$，输出层神经元 o_1 的输出 $a_{k=1} = 0.751365$；
- 输出层神经元 o_1 的输入 $in_{o_1} = \sum_j w_{j,k=1}a_j = 1.10596$，sigmoid 函数在 in_{o_1} 的导数 $g'(in_{o_1}) = \text{sigmoid}'(1.10596) = \text{sigmoid}(1.10596) \times (1 - \text{sigmoid}(1.10596)) = 0.186816$；
- $w_{j=1,k=1} = w_5 = 0.4$；
- 隐层神经元 h_1 的输入 $in_{h_1} = \sum_i w_{i,j=1}a_i = 0.3775$，sigmoid 函数在 in_{h_1} 的导数 $g'(in_{h_1}) = 0.2413$；
- 输入 i_1 的值 $a_{i=1} = 0.05$。

因此有

$$\frac{\partial Loss_{o_1}}{\partial w_1} = -2(0.01 - 0.751365) \times 0.186816 \times 0.4 \times 0.2413 \times 0.05 \approx 0.00133679$$

(b) 要计算 $\frac{\partial Loss_{o_2}}{\partial w_8}$，即是要计算 $\frac{\partial Loss_{k=2}}{\partial w_{j=2,k=2}}$。从课件数据计算式中各项，可得

- $y_2 = 0.99$，输出层神经元 o_2 的输出 $a_{k=2} = 0.755523$；
- 输出层神经元 o_2 的输入 $in_{o_2} = \sum_j w_{j,k=2}a_j = 0.985700$，tanh 函数在 in_{o_2} 的导数 $g'(in_{o_2}) = 1 - g^2(in_{o_2}) = 1 - \tanh^2(0.9857) = 2\text{sigmoid}(1.9714) - 1 = 0.429186$；
- 隐层神经元 h_2 的输出 $a_{j=2} = g(\sum_i w_{i,j=2}a_i) = \tanh(0.25 \times 0.05 + 0.3 \times 0.1 + 0.35) = 2\text{sigmoid}(2 \times 0.3925) - 1 = 0.373513$。

因此有

$$\frac{\partial Loss_{o_2}}{\partial w_8} = -2(0.99 - 0.755523) \times 0.429186 \times 0.373513 \approx -0.0751764$$

4. Q4 感谢邓义圣同学的答案

解:(a)因为图像为 5*5*3,卷积核大小为 3*3*3,且卷积核只有一个,则(5-3)/2+1=2,则输出大小为 2*2*1.设四个值为 a0,a1,a2,a3

a0=(1*1+1*0+2*(-1)+1*1+0*0+0*(-1)+0*1+1*0+0*(-1))+(0*1+2*1+1*1+1*0+0*0+2*0+1*(-1)+0*(-1)+2*(-1))+(1*1+1*0+0*1+1*0+1*(-1)+0*0+0*1+2*0+1*1)+1=2

对于 a1,a2,a3,有类似的结果,由于权重共享,只需要将每个乘号前一项改为新的值即可.

a1=0

a2=6

a3=3

故最终结果如下:

2	0
6	3

(b)在 2*2 的最大池化中,输出结果为 max(a0,a1,a2,a2),a0,a1,a2,a2 为矩阵中的 4 个值.

对于最左上角的 2*2 矩阵,输出为 max(6,0,-7,4)=6

同理,对于剩下的输出结果依次为 16,10,15,14,14,4,19,0

6	16	10
15	14	14
4	19	0

(c)因为卷积核个数为 16,卷积核大小为 6*6*3,输入大小为 100*100*3,步长为 3,填充位为 1,则经过卷积层后输出大小为((100+1*2-6)/3+1)*((100+1*2-6)/3+1)*16

即 33*33*16

经过池化层后 W=(33-8)/5+1=6

H=(33-8)/5+1=6

D 不变,仍为 16.

则池化层后输出为 6*6*16

5.

① 选择 healthy

```
Q[healthy, party] = 0.7*(10 + 0.9 * 0) + 0.3*(10 + 0.9 * 0) = 10
Q[healthy, relax] = 0.95*(7 + 0.9 * 0) + 0.05*(7 + 0.9 * 0) = 7
V[healthy] = max(10, 7) = 10
```

Q[s,a]	party	relax
healthy	10	7
sick	0	0

	healthy	sick
V[s]	10	0

② 选择 sick

```
Q[sick, party] = 0.1*(2 + 0.9 * 10) + 0.9*(2 + 0.9 * 0) = 2.9
Q[sick, relax] = 0.5*(0 + 0.9 * 10) + 0.5*(0 + 0.9 * 0) = 4.5
V[sick] = max(2.9, 4.5) = 4.5
```

Q[s,a]	party	relax
healthy	10	7
sick	2.9	4.5

	healthy	sick
V[s]	10	4.5

③ 选择 healthy

```
Q[healthy, party] = 0.7*(10 + 0.9 * 10) + 0.3*(10 + 0.9 * 4.5) = 17.515
Q[healthy, relax] = 0.95*(7 + 0.9 * 10) + 0.05*(7 + 0.9 * 4.5) = 15.7525
V[healthy] = max(17.515, 15.7525) = 17.515
```

Q[s,a]	party	relax
healthy	17.515	15.7525
sick	2.9	4.5

	healthy	sick
V[s]	17.515	4.5

感谢吴莹菲同学的答案