AI(Fall 2020) - Assignment 4

Machine Learning

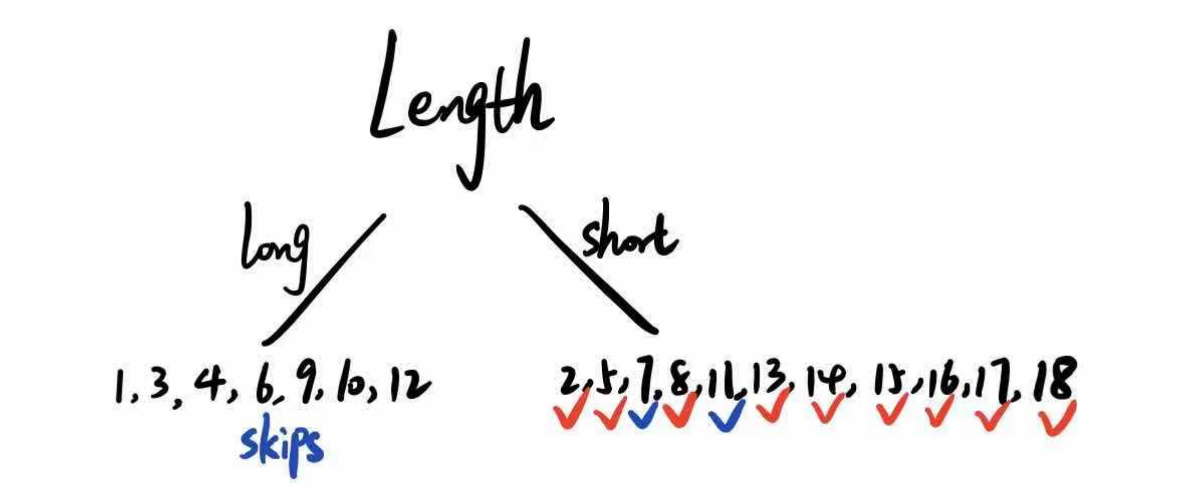
1. Consider the following data. The DECISION-TREE-LEARNING algorithm will first select the attribute Length to split on. Finish building the decision tree, and show the computations.

Example	Author	Thread	Length	Where Read	User Action
e1	known	new	long	home	skips
e2	unknown	new	short	work	reads
e3	unknown	follow Up	long	work	skips
e4	known	follow Up	long	home	skips
e5	known	new	short	home	reads
e6	known	follow Up	long	work	skips
e7	unknown	follow Up	short	work	skips
e8	unknown	new	short	work	reads
e9	known	follow Up	long	home	skips
e10	known	new	long	work	skips
e11	unknown	follow Up	short	home	skips
e12	known	new	long	work	skips
e13	known	follow Up	short	home	reads
e14	known	new	short	work	reads
e15	known	new	short	home	reads
e16	known	follow Up	short	work	reads
e17	known	new	short	home	reads
e18	unknown	new	short	work	reads

Answer:

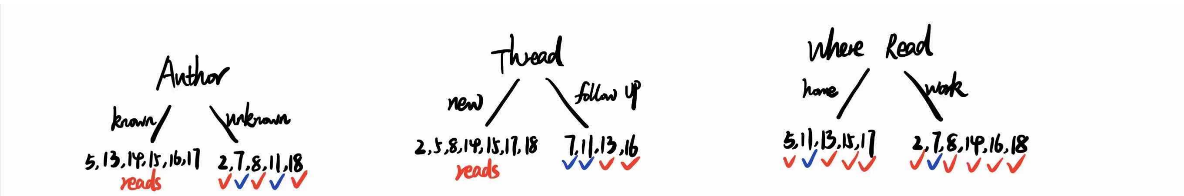
注：蓝色代表skips，红色代表reads

首先按照`Length`分裂得到决策树如下：



可以看到左子树均为`skips`，故不用再继续分裂；右子树则需再进行分裂。

按照剩下的三个属性：`Author`，`Thread`，`Where Read`分别进行分裂，结果如下：



接下来需要计算信息增益，哪个属性对应的信息增益最大就选择哪个属性进一步分裂：

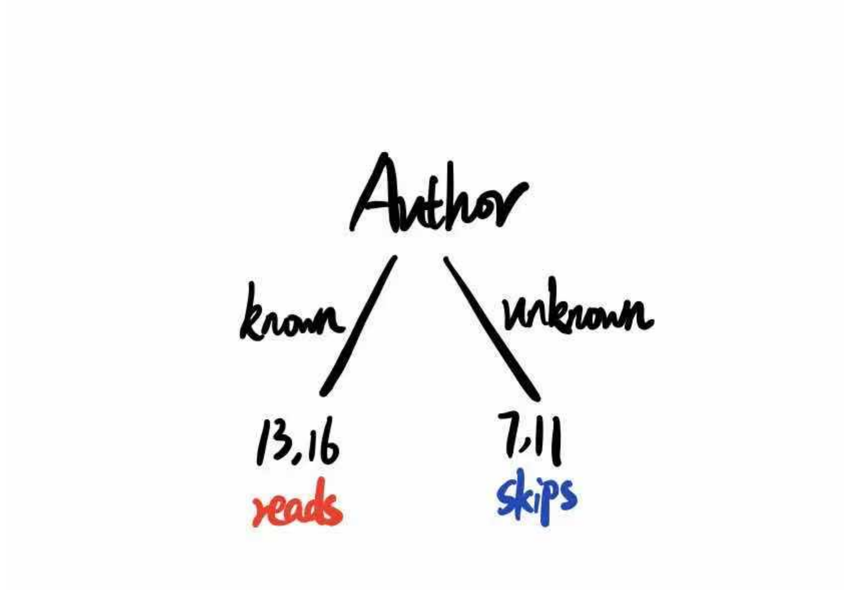
$$
\begin{aligned}
&Ent(D) = B(\tfrac{2}{11}) = -(\tfrac{2}{11}log_2\tfrac{2}{11} + \tfrac{9}{11}log_2\tfrac{9}{11}) = 0.684038 \\
&Gain(D, Author) = Ent(D) - (\tfrac{6}{11}B(\tfrac{0}{6}) + \tfrac{5}{11}B(\tfrac{2}{5})) \\
&= 0.684038 - (-\tfrac{6}{11}(\tfrac{6}{6}log_2\tfrac{6}{6} + \tfrac{0}{6}log_2\tfrac{0}{6}) - \tfrac{5}{11}(\tfrac{3}{5}log_2\tfrac{3}{5} + \tfrac{2}{5}log_2\tfrac{2}{5})) \\
&= 0.242697 \\
&Gain(D, Thread) = Ent(D) - (\tfrac{7}{11}B(\tfrac{0}{7}) + \tfrac{4}{11}B(\tfrac{2}{4})) \\
&= 0.684038 - (-\tfrac{7}{11}(\tfrac{7}{7}log_2\tfrac{7}{7} + \tfrac{0}{7}log_2\tfrac{0}{7}) - \tfrac{4}{11}(\tfrac{2}{4}log_2\tfrac{2}{4} + \tfrac{2}{4}log_2\tfrac{2}{4})) \\
&= 0.320402 \\
&Gain(D, WhereRead) = Ent(D) - (\tfrac{5}{11}B(\tfrac{1}{5}) + \tfrac{6}{11}B(\tfrac{1}{6})) \\
&= 0.684038 - (-\tfrac{5}{11}(\tfrac{4}{5}log_2\tfrac{4}{5} + \tfrac{1}{5}log_2\tfrac{1}{5}) - \tfrac{6}{11}(\tfrac{1}{6}log_2\tfrac{1}{6} + \tfrac{5}{6}log_2\tfrac{5}{6})) \\
&= 0.001331
\end{aligned}
$$

可以看到最大的为$Gain(D, Thread)$,因此按$Thread$进行分裂,可以看到左子树均为$reads$不用再分裂;右子树按剩下的两个属性进行分裂：

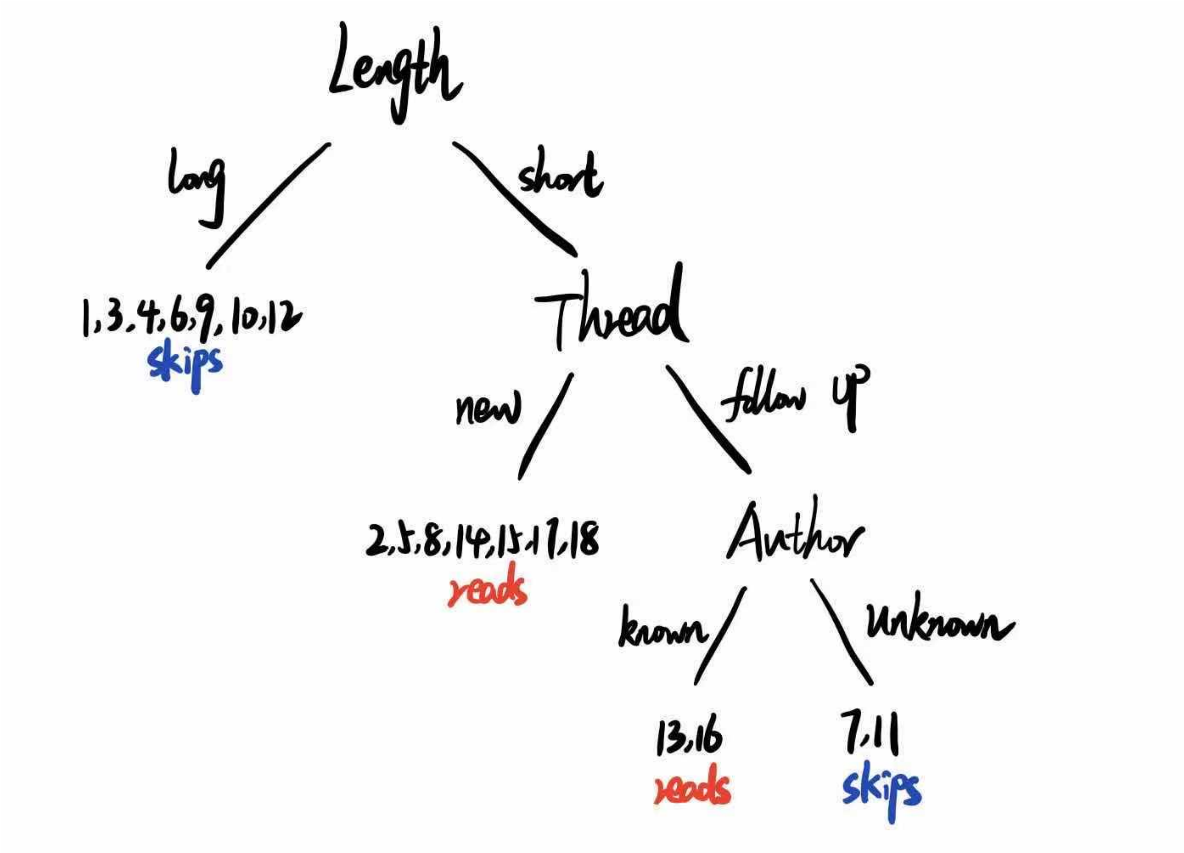
重复上述步骤：

$$
\begin{aligned}
&Ent(D) = B(\tfrac{2}{4}) = -(\tfrac{2}{4}log_2\tfrac{2}{4} + \tfrac{2}{4}log_2\tfrac{2}{4}) = 1 \\
&Gain(D, Author) = Ent(D) - (\tfrac{2}{4}B(\tfrac{0}{2}) + \tfrac{2}{4}B(\tfrac{2}{2})) \\
&= 1 - (-\tfrac{2}{4}(\tfrac{2}{2}log_2\tfrac{2}{2} + \tfrac{0}{2}log_2\tfrac{0}{2}) - \tfrac{2}{4}(\tfrac{0}{2}log_2\tfrac{0}{2} + \tfrac{2}{2}log_2\tfrac{2}{2})) \\
&= 1 \\
&Gain(D, WhereRead) = Ent(D) - (\tfrac{2}{4}B(\tfrac{1}{2}) + \tfrac{2}{4}B(\tfrac{1}{2})) \\
&= 1 - (-\tfrac{2}{4}(\tfrac{1}{2}log_2\tfrac{1}{2} + \tfrac{1}{2}log_2\tfrac{1}{2}) - \tfrac{2}{4}(\tfrac{1}{2}log_2\tfrac{1}{2} + \tfrac{1}{2}log_2\tfrac{1}{2})) \\
&= 0
\end{aligned}
$$

可以看到最大的为$Gain(D, Author)$,故按$Author$分裂,分裂之后如下：



可以看到左右子树内的元素均属于同一类别,至此不用再分裂了.故最终得到的决策树如下:



2. Consider the candy example from the lecture. Assume that the prior distribution over $h_1, \ldots, h_5$ is given by $\langle 0.1, 0.2, 0.4, 0.2, 0.1 \rangle$. Suppose that the first 5 candies taste lime, cherry, cherry, lime, and lime. Make predictions for the 6th candy using Bayesian, MAP and ML learning, respectively. Show the computations done to make the predictions.

Answer:

原题如下:

- Hypothesis H: probabilistic theory of the world
  - h_1: 100% cherry
  - h_2: 75% cherry + 25% lime
  - h_3: 50% cherry + 50% lime
  - h_4: 25% cherry + 75% lime
  - h_5: 100% lime

记$d =< lime, cherry, cherry, lime, lime >$，相关信息的表格如下:

hypothesis	h_1	h_2	h_3	h_4	h_5
$P(lime\|h_i)$	0	0.25	0.5	0.75	1
$P(cherry\|h_i)$	1	0.75	0.5	0.25	0
$P(h_i)$	0.1	0.2	0.4	0.2	0.1
$P(d\|h_i)$	0	$(0.25)^3 \times (0.75)^2 = \frac{9}{1024}$	同理可得为 $\frac{1}{32}$	同理可得为 $\frac{27}{1024}$	0

1. `Bayesian`:

$$P(d) = \Sigma_i P(d|h_i)P(h_i) = 0.2 \times \frac{9}{1024} + 0.4 \times \frac{1}{32} + 0.2 \times \frac{27}{1024} = 0.01953125$$

$$P(lime|d) = \frac{\Sigma_i P(lime|h_i)P(d|h_i)P(h_i)}{P(d)} = \frac{0.25 \times \frac{9}{1024} \times 0.2 + 0.5 \times \frac{1}{32} \times 0.4 + 0.75 \times \frac{27}{1024} \times 0.2}{0.01953125} = 0.545$$

$$P(cherry|d) = \frac{\Sigma_i P(cherry|h_i)P(d|h_i)P(h_i)}{P(d)} = \frac{0.75 \times \frac{9}{1024} \times 0.2 + 0.5 \times \frac{1}{32} \times 0.4 + 0.25 \times \frac{27}{1024} \times 0.2}{0.01953125} = 0.455$$

因为$P(lime|d) > P(cherry|d)$,故预测第6次是$lime$

2. `MAP`:

$$h_{MAP} = \underset{h_i}{argmax} P(h_i|d) = \underset{h_i}{argmax} P(d|h_i)P(h_i)$$

又因为

$$P(d|h_1)P(h_1) = 0$$
$$P(d|h_2)P(h_2) = 0.2 \times \frac{9}{1024} = 0.0017578125$$
$$P(d|h_3)P(h_3) = 0.4 \times \frac{1}{32} = 0.0125$$
$$P(d|h_4)P(h_4) = 0.2 \times \frac{27}{1024} = 0.0052734375$$
$$P(d|h_5)P(h_5) = 0$$

故$h_{MAP} = h_3$

因为$P(lime|h_3) = 0.5, P(cherry|h_3) = 0.5$

第六次为$lime$和$cherry$的预测概率相等.

3. `ML`:

$$h_{ML} = \underset{h_i}{argmax} P(d|h_i)$$

根据表格可以知道当h_i取h_3时,$P(d|h_i)$最大

故$P(lime|h_{ML}) = P(lime|h_3) = 0.5, P(cherry|h_{ML}) = P(cherry|h_3) = 0.5$

第六次为$lime$和$cherry$的预测概率相等.

3. Consider the Boolean function E = (A XOR B) AND (C XOR D). Construct its truth table, and then remove the line for the input $A = 1, B = 1, C = 1, D = 1$. Use Naive Bayes classification to make prediction for this input. Show the computations.

Answer:

对于A, B, C, D的取值，共有$2^4 = 16$种取法，不考虑题目待预测的取值则有15种取法，首先要将这15种取法的真值表得到作为样本：

已知样本	A	B	C	D	E
1	0	0	0	0	0
2	0	0	0	1	0
3	0	0	1	0	0
4	0	0	1	1	0
5	0	1	0	0	0
6	0	1	0	1	1
7	0	1	1	0	1
8	0	1	1	1	0
9	1	0	0	0	0
10	1	0	0	1	1
11	1	0	1	0	1
12	1	0	1	1	0
13	1	1	0	0	0
14	1	1	0	1	0
15	1	1	1	0	0

然后计算概率:

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

由于二式分母均为$P(A,B,C,D)$，因此可以不计算该值，直接比较分子大小即可：

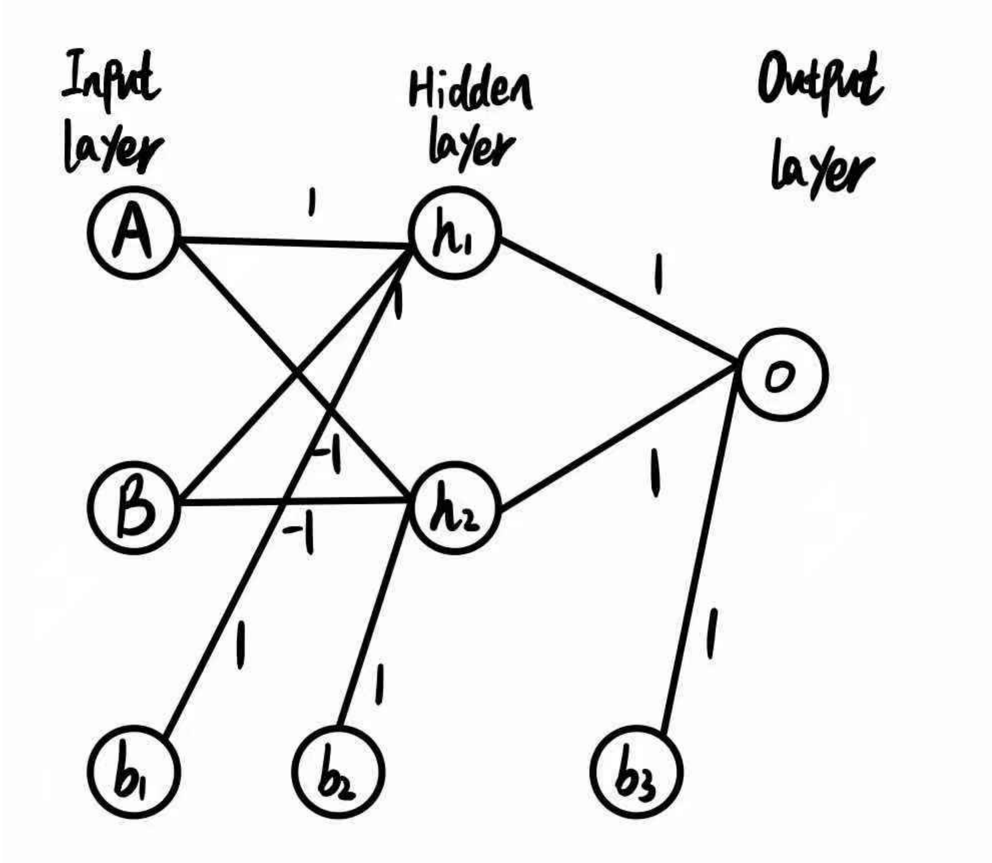
$$P(A|E)P(B|E)P(C|E)P(D|E)P(E)=\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{4}{15}=\frac{1}{60}$$
$$P(A|\neg E)P(B|\neg E)P(C|\neg E)P(D|\neg E)P(\neg E)=\frac{5}{11}\times\frac{5}{11}\times\frac{5}{11}\times\frac{5}{11}\times\frac{11}{15}=\frac{125}{3993}$$

因为$\frac{1}{60}<\frac{125}{3993}$,故$P(E|A,B,C,D)<P(\neg E|A,B,C,D)$,故判别$E=0.$

4. Construct a neural network that computes the XOR function of two inputs.

Answer:

构建如下的神经网络（边上的数字为权重）：



其中$b_1 = -0.5, b_2 = 1.5, b_3 = -1.5$，激活函数$g(x)$定义为：

$$g(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

故根据`Forward pass`可以计算四种情况下的输出结果：

$$\begin{aligned}
&out(0,0) = g(g(0\times 1 + 0\times 1 - 0.5) * 1 + g(0\times(-1) + 0\times(-1) + 1.5) * 1 - 1.5) \\
&= g(0 + 1 - 1.5) = g(-0.5) = 0 \\
&out(0,1) = g(g(0\times 1 + 1\times 1 - 0.5) * 1 + g(0\times(-1) + 1\times(-1) + 1.5) * 1 - 1.5) \\
&= g(1 + 1 - 1.5) = g(0.5) = 1 \\
&out(1,0) = g(g(1\times 1 + 0\times 1 - 0.5) * 1 + g(1\times(-1) + 0\times(-1) + 1.5) * 1 - 1.5) \\
&= g(1 + 1 - 1.5) = g(0.5) = 1 \\
&out(1,1) = g(g(1\times 1 + 1\times 1 - 0.5) * 1 + g(1\times(-1) + 1\times(-1) + 1.5) * 1 - 1.5) \\
&= g(1 + 0 - 1.5) = g(-0.5) = 0
\end{aligned}$$

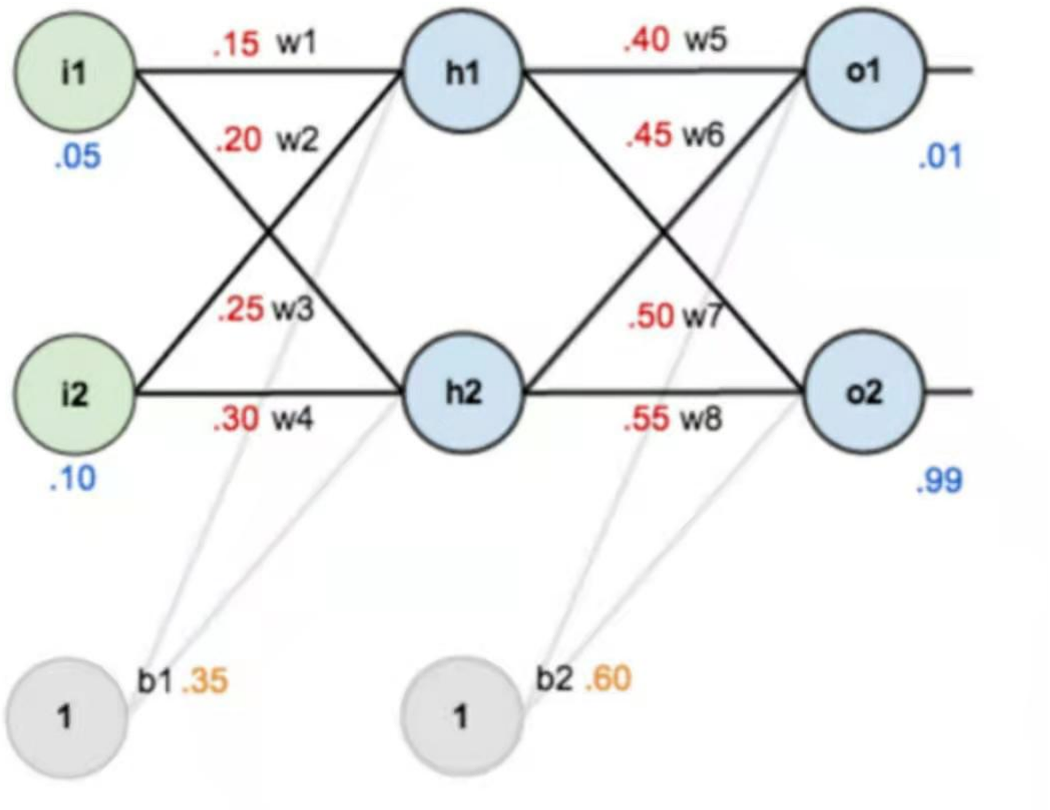
可以看到输出与两个输入间的结果满足异或关系。

5. Consider the neural net on Page 32 of the course slides for neural nets.

- Suppose we use the sigmoid function as the activate function. Compute $\partial Loss_{o_1}/\partial w_1$.
- Suppose we use the tanh function as the activate function. Compute $\partial Loss_{o_2}/\partial w_4$. Note that $tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}) = 2g(2x) - 1$, and $tanh'(x) = 1 - tanh^2(x)$.

Answer:

1.课件给出的神经网络如下：



课件给出的损失函数为$Loss_k=(y_k-a_k)^2$，计算如下：

$in_{h1}=w_1i_1+w_2i_2+b_1=0.15\times0.05+0.20\times0.10+0.35=0.3775$
$out_{h1}=g(in_{h1})=\frac{1}{1+e^{-0.3775}}=0.593269992$
$in_{h2}=w_3i_1+w_4i_2+b_1=0.25\times0.05+0.30\times0.10+0.35=0.3925$
$out_{h2}=g(in_{h2})=\frac{1}{1+e^{-0.3925}}=0.596884378$
$in_{o1}=w_5out_{h1}+w_6out_{h2}+b_2=0.40\times0.593269992+0.45\times0.596884378+0.60=1.105905967$
$out_{o1}=g(in_{o1})=\frac{1}{1+e^{-1.105905967}}=0.751365070$
$in_{o2}=w_7out_{h1}+w_8out_{h2}+b_2=0.50\times0.593269992+0.55\times0.596884378+0.60=1.224921404$
$out_{o2}=g(in_{o2})=\frac{1}{1+e^{-1.224921404}}=0.772928465$
$Loss_{o1}=(target_{o1}-out_{o1})^2=0.549622167$
$Loss_{o2}=(target_{o2}-out_{o2})^2=0.047121354$
$\frac{\partial Loss_{o1}}{\partial w_1}=-2(target_{o1}-out_{o1})\frac{\partial out_{o1}}{\partial w_1}=-2(target_{o1}-out_{o1})(out_{o1}(1-out_{o1}))\frac{\partial in_{o1}}{\partial w_1}$
$=-2(target_{o1}-out_{o1})(out_{o1}(1-out_{o1}))w_5\frac{\partial out_{h1}}{\partial w_1}$
$=-2(target_{o1}-out_{o1})(out_{o1}(1-out_{o1}))w_5(out_{h1}(1-out_{h1}))\frac{\partial in_{h1}}{\partial w_1}$
$=-2(target_{o1}-out_{o1})(out_{o1}(1-out_{o1}))w_5(out_{h1}(1-out_{h1}))i_1$
$=-2\times(0.01-0.751365070)\times(0.751365070)\times(1-0.751365070)\times0.40\times(0.593269992)\times$
$(1-0.593269992)\times0.05=0.001336792$

2.过程类似，只不过将激活函数由$\frac{1}{1+e^{-x}}$换成了$tanh(x)$:

$in_{h1} = w_1 i_1 + w_2 i_2 + b_1 = 0.15 \times 0.05 + 0.20 \times 0.10 + 0.35 = 0.3775$
$out_{h1} = tanh(in_{h1}) = 2g(2in_{h1}) - 1 = 0.360534393$
$in_{h2} = w_3 i_1 + w_4 i_2 + b_1 = 0.25 \times 0.05 + 0.30 \times 0.10 + 0.35 = 0.3925$
$out_{h2} = tanh(in_{h2}) = 2g(2in_{h2}) - 1 = 0.373513453$
$in_{o1} = w_5 out_{h1} + w_6 out_{h2} + b_2 = 0.40 \times 0.360534393 + 0.45 \times 0.373513453 + 0.60 = 0.912294811$
$out_{o1} = tanh(in_{o1}) = 2g(2in_{o1}) - 1 = 0.722231868$
$in_{o2} = w_7 out_{h1} + w_8 out_{h2} + b_2 = 0.50 \times 0.360534393 + 0.55 \times 0.373513453 + 0.60 = 0.985699596$
$out_{o2} = tanh(in_{o2}) = 2g(2in_{o2}) - 1 = 0.755522642$
$Loss_{o1} = (target_{o1} - out_{o1})^2 = 0.507274234$
$Loss_{o2} = (target_{o2} - out_{o2})^2 = 0.054979631$
$\frac{\partial Loss_{o2}}{\partial w_4} = -2(target_{o2} - out_{o2}) \frac{\partial out_{o2}}{\partial w_4} = -2(target_{o2} - out_{o2})(1 - out_{o2}^2) \frac{\partial in_{o2}}{\partial w_4}$
$= -2(target_{o2} - out_{o2})(1 - out_{o2}^2) w_8 \frac{\partial out_{h2}}{\partial w_4}$
$= -2(target_{o2} - out_{o2})(1 - out_{o2}^2) w_8 (1 - out_{h2}^2) \frac{\partial in_{h2}}{\partial w_4}$
$= -2(target_{o2} - out_{o2})(1 - out_{o2}^2) w_8 (1 - out_{h2}^2) i_2$
$= -2 \times (0.99 - 0.755522642) \times (1 - 0.755522642^2) \times 0.55 \times (1 - 0.373513453^2) \times 0.10$
$= -0.009525403$