

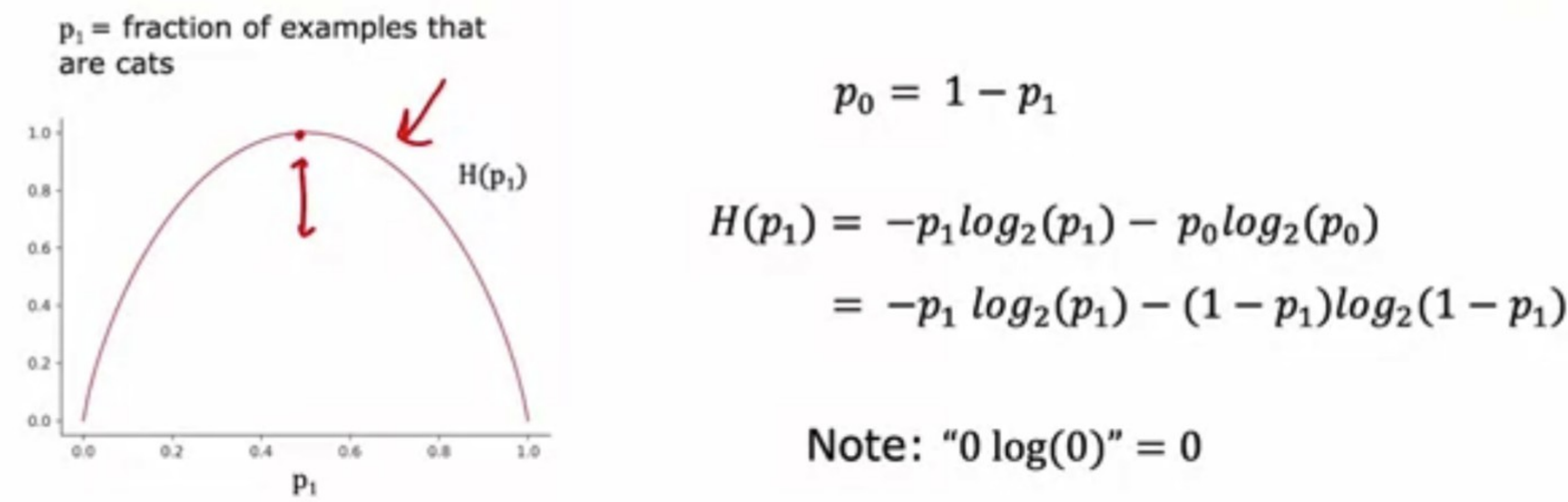
✔ 恭喜！您通过了！

获得的成绩 100% 最新提交作业的评分 100% 通过条件 80% 或更高

转到下一个课程内容

① 您正在查看此评估的自动翻译版本
您可以根据需要切换回去，以源语言查看这些内容。即使您更改语言，也不会丢失任何进度。
[显示英语（English）版本](#) 取消 ×

1. Entropy as a measure of impurity 1 / 1个点



回想一下，熵在讲座中的定义是 $H(p_1) = -p_1 \log_2(p_1) - p_0 \log_2(p_0)$ ，其中 p_1 是正面例子的比例， p_0 是反面例子的比例。

在一棵决策树的给定节点上，10 个例子中有 6 个是猫，4 个不是猫。哪个表达式可以计算这 10 种动物的熵 $H(p_1)$ ？

- ☒ $-(0.6)\log_2(0.6) - (0.4)\log_2(0.4)$
- ☐ $(0.6)\log_2(0.6) + (0.4)\log_2(0.4)$
- ☐ $-(0.6)\log_2(0.6) - (1 - 0.4)\log_2(1 - 0.4)$
- ☐ $(0.6)\log_2(0.6) + (1 - 0.4)\log_2(1 - 0.4)$

✔ 正确
正确。表达式是 $-(p_1)\log_2(p_1) - (p_0)\log_2(p_0)$

2. Information gain 1 / 1个点

Information gain

$$= H(p_1^{root}) - \left(w^{left} H(p_1^{left}) + w^{right} H(p_1^{right}) \right)$$

回顾信息的定义如下：











$$H(p_1^{root}) - \left(w^{left} H(p_1^{left}) + w^{right} H(p_1^{right}) \right)$$

在拆分之前，由 5 只猫和 5 只非猫组成的一组的熵为 $H(5/10)$ 。在对某一特征进行拆分后，一组 7 种动物（其中 4 种是猫）的熵为 $H(4/7)$ 。另一组 3 种动物（其中 1 种是猫）的熵为 $H(1/3)$ 。信息增益的表达式是什么？

- ☐ $H(0.5) - (H(4/7) + H(1/3))$
- ☐ $H(0.5) - \left(\frac{4}{7} * H(4/7) + \frac{4}{7} * H(1/3) \right)$
- ☐ $H(0.5) - (7 * H(4/7) + 3 * H(1/3))$
- ☒ $H(0.5) - \left(\frac{7}{10} H(4/7) + \frac{3}{10} H(1/3) \right)$

✔ 正确
正确。一般表达式为 $H(p_1^{root}) - \left(w^{left} H(p_1^{left}) + w^{right} H(p_1^{right}) \right)$

One hot encoding

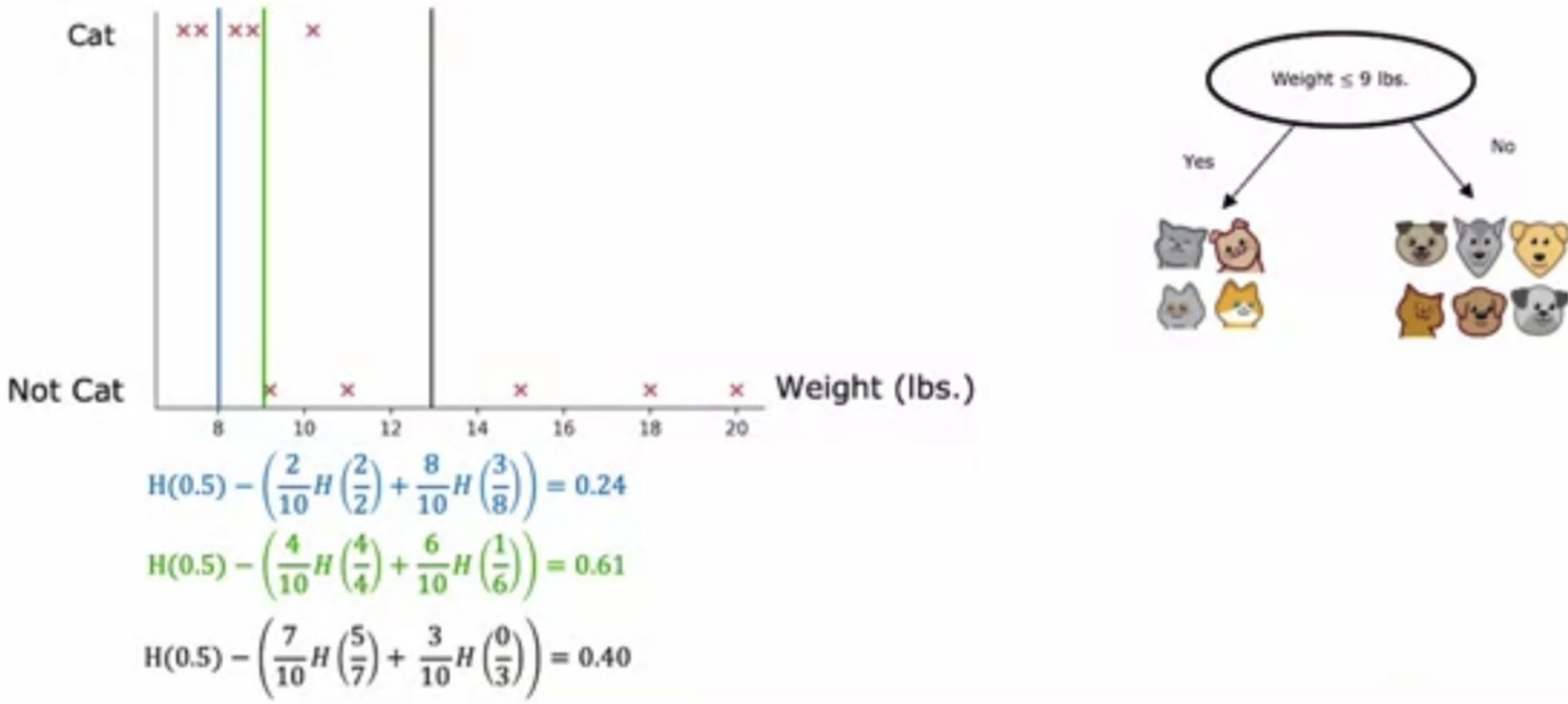
Ear-shape	Pointy ears	Floppy ears	Oval ears	Face shape	Whiskers	Cat
 Pointy	1	0	0	Round	Present	1
 Oval	0	0	1	Not round	Present	1
 Oval	0	0	1	Round	Absent	0
 Pointy	1	0	0	Not round	Present	0
 Oval	0	0	1	Round	Present	1
 Pointy	1	0	0	Round	Absent	1
 Floppy	0	1	0	Not round	Absent	0
 Oval	0	0	1	Round	Absent	1
 Floppy	0	1	0	Round	Absent	0
 Floppy	0	1	0	Round	Absent	0

要表示耳朵形状的 3 种可能值，可以定义 3 种耳朵形状特征：尖耳朵、软耳朵、椭圆形耳朵。对于耳朵不是尖尖的、不是软软的，而是椭圆形的动物，如何用特征向量来表示这一信息？

- ☒ [0, 0, 1]
- ☐ [1, 1, 0]
- ☐ [1,0,0]
- ☐ [0, 1, 0]

☒ 正确
是的！ 0 表示不存在该特征（不尖、不软），1 表示存在该特征（椭圆形）。

Splitting on a continuous variable



对于连续值特征（如动物体重），数据集中有 10 只动物。根据讲座内容，建议如何为该特征找到最佳分割？

- ☐ 使用梯度下降法找出信息增益最大的分割阈值。
- ☐ 使用单次编码将特征转化为由 0 和 1 组成的离散特征向量，然后应用我们讨论过的离散特征算法。
- ☒ 选择 10 个例子中的 9 个中间点作为可能的分割点，并找出信息增益最高的分割点。
- ☐ 尝试每一个间隔固定的数值（如 8、8.5、9、9.5、10 等），找出信息增益最高的分。

☒ 正确
正确。这就是讲座中提出的建议。

哪些是决定停止拆分的常用标准？(选择两项)。

- ☐ 当额外拆分带来的信息增益过大时
- ☒ 当树木达到最大深度时

☒ 正确
是的！

- ☒ 当节点中的示例数量低于阈值时

☒ 正确
是的！

- ☐ 当一个节点 50% 属于一个类别，50% 属于另一个类别时（熵的最高值）