

Python 3 玩儿转机器学习

讲师：liuyubobobo

版权所有 侵权必究
liuyubobobo

慕课网《Python3机器学习》

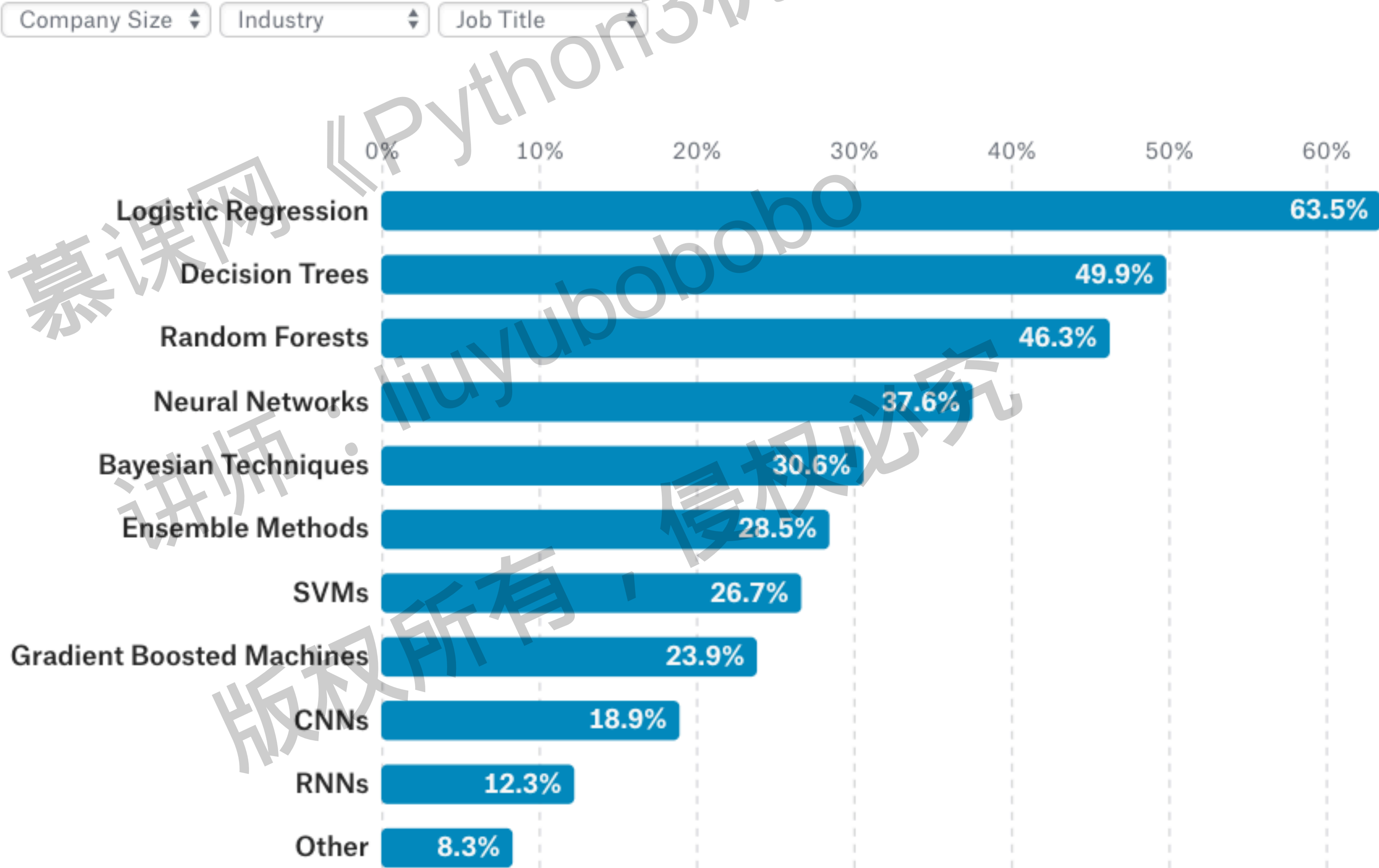
逻辑回归

Logistic Regression

讲师: julyuboo
版权所有, 侵权必究

What data science methods are used at work?

Logistic regression is the most commonly reported data science method used at work for all industries *except* **Military and Security** where Neural Networks are used slightly more frequently.



慕课网《Python3机器学习》

什么是逻辑回归

讲师：liuyuboboo

版权所有，侵权必究

逻辑回归 Logistic Regression

逻辑回归：解决分类问题

回归问题怎么解决分类问题？

将样本的特征和样本发生的概率联系起来，概率是一个数

逻辑回归 Logistic Regression

$$\hat{y} = f(x)$$

$$\hat{p} = f(x) \quad \hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} \leq 0.5 \end{cases}$$

逻辑回归既可以看做是回归算法，也可以看做是分类算法

通常作为分类算法用，只可以解决二分类问题

逻辑回归 Logistic Regression

$$\hat{y} = f(x) \rightarrow \hat{y} = \theta^T \cdot x_b$$

值域 (-infinity, +infinity)

概率的值域为 [0, 1]

$$\hat{p} = \sigma(\theta^T \cdot x_b)$$

Sigmoid 函数

$$\hat{p} = \sigma(\theta^T \cdot x_b)$$

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

慕课网《Python3机器学习》

实践：sigmoid函数

讲师：liuyubobobo

版权所有，侵权必究

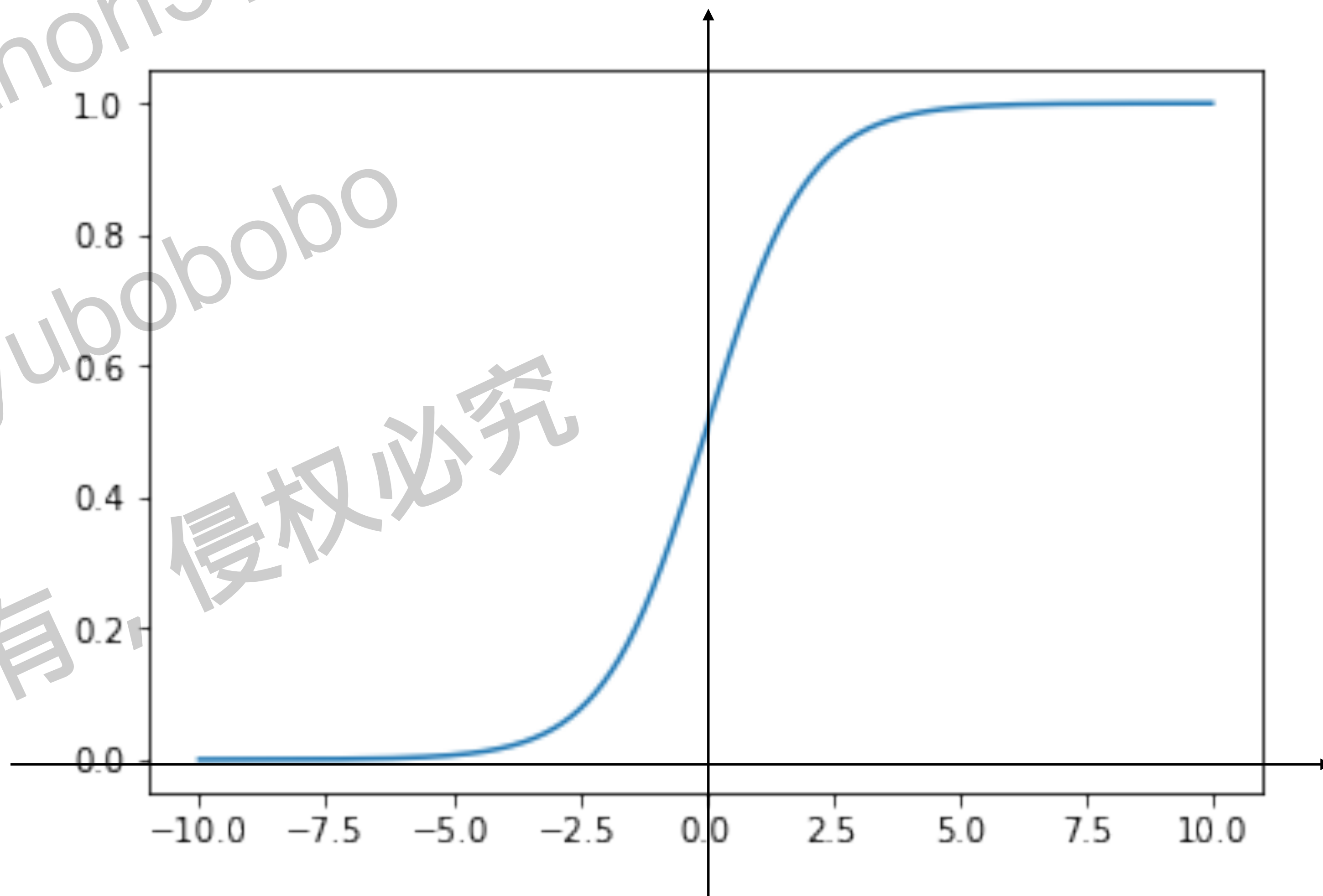
Sigmoid 函数

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

值域 (0, 1)

$t > 0$ 时, $p > 0.5$

$t < 0$ 时, $p < 0.5$



逻辑回归 Logistic Regression

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} \leq 0.5 \end{cases}$$

逻辑回归 Logistic Regression

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} \leq 0.5 \end{cases}$$

问题：

对于给定的样本数据集X,y,

我们如何找到参数theta,

使得用这样的方式,

可以最大程度获得样本数据集X

对应的分类输出y?

逻辑回归中损失函数的定义

讲师：liuyuboboo
版权所有，侵权必究

逻辑回归 Logistic Regression

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} \leq 0.5 \end{cases}$$

逻辑回归 Logistic Regression

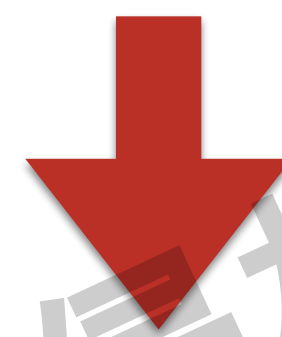
$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} \leq 0.5 \end{cases}$$

$$\text{cost} = \begin{cases} \text{如果 } y=1, p \text{ 越小, cost 越大} \\ \text{如果 } y=0, p \text{ 越大, cost 越大} \end{cases}$$

逻辑回归 Logistic Regression

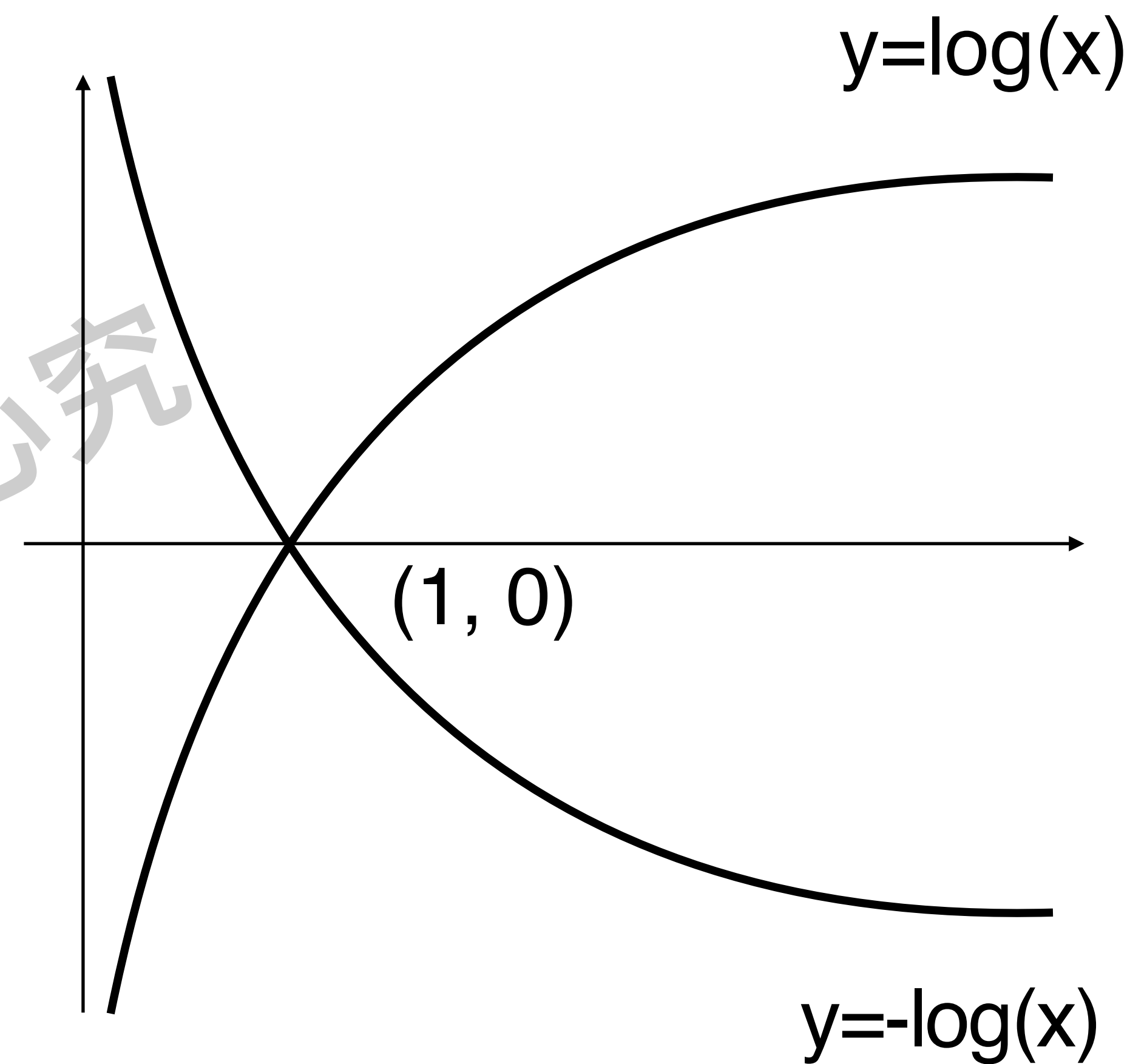
$$\text{cost} = \begin{cases} \text{如果 } y=1, p \text{ 越小, cost 越大} \\ \text{如果 } y=0, p \text{ 越大, cost 越大} \end{cases}$$



$$\text{cost} = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1 - \hat{p}) & \text{if } y = 0 \end{cases}$$

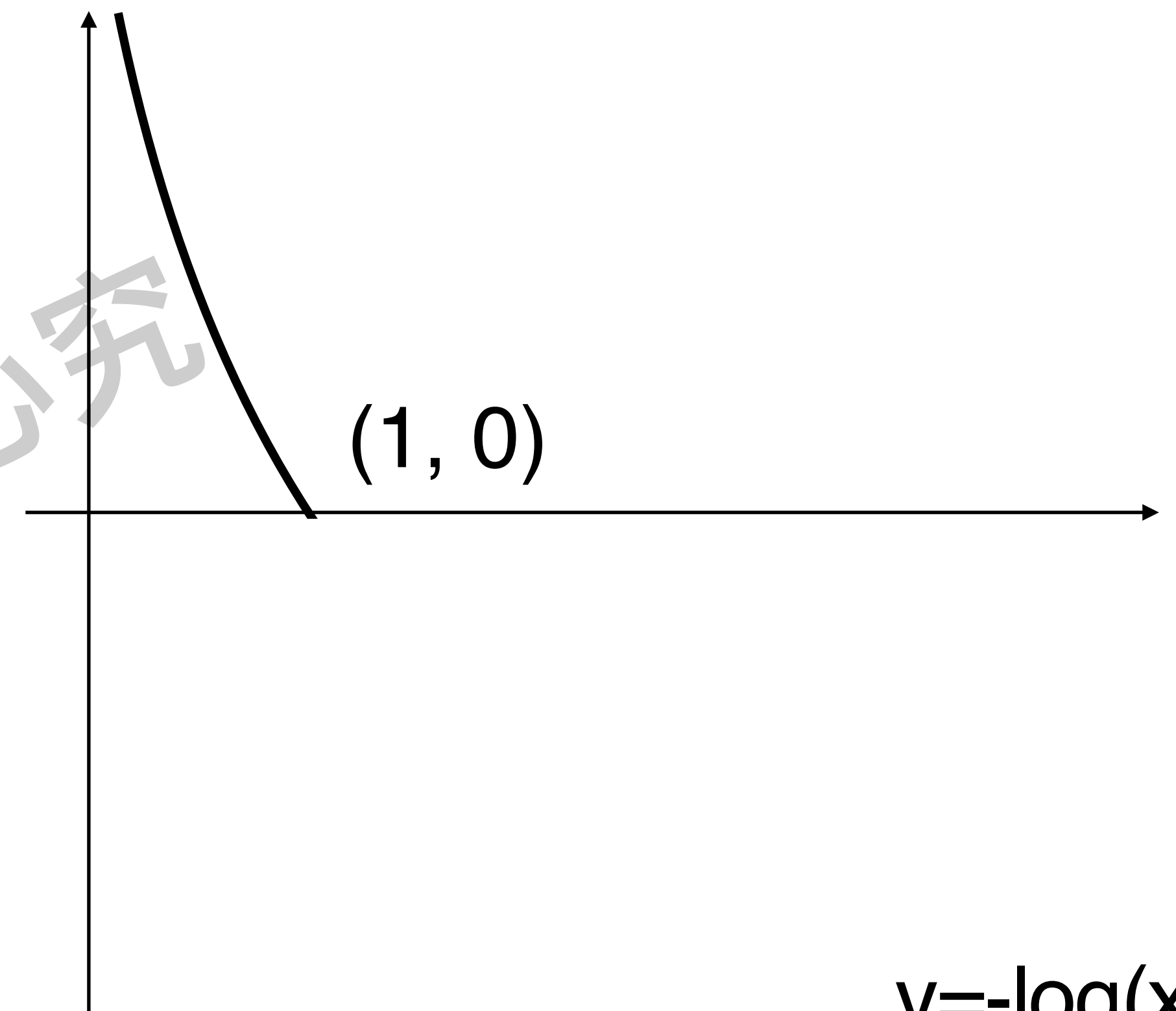
逻辑回归 Logistic Regression

$$\text{cost} = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1 - \hat{p}) & \text{if } y = 0 \end{cases}$$



逻辑回归 Logistic Regression

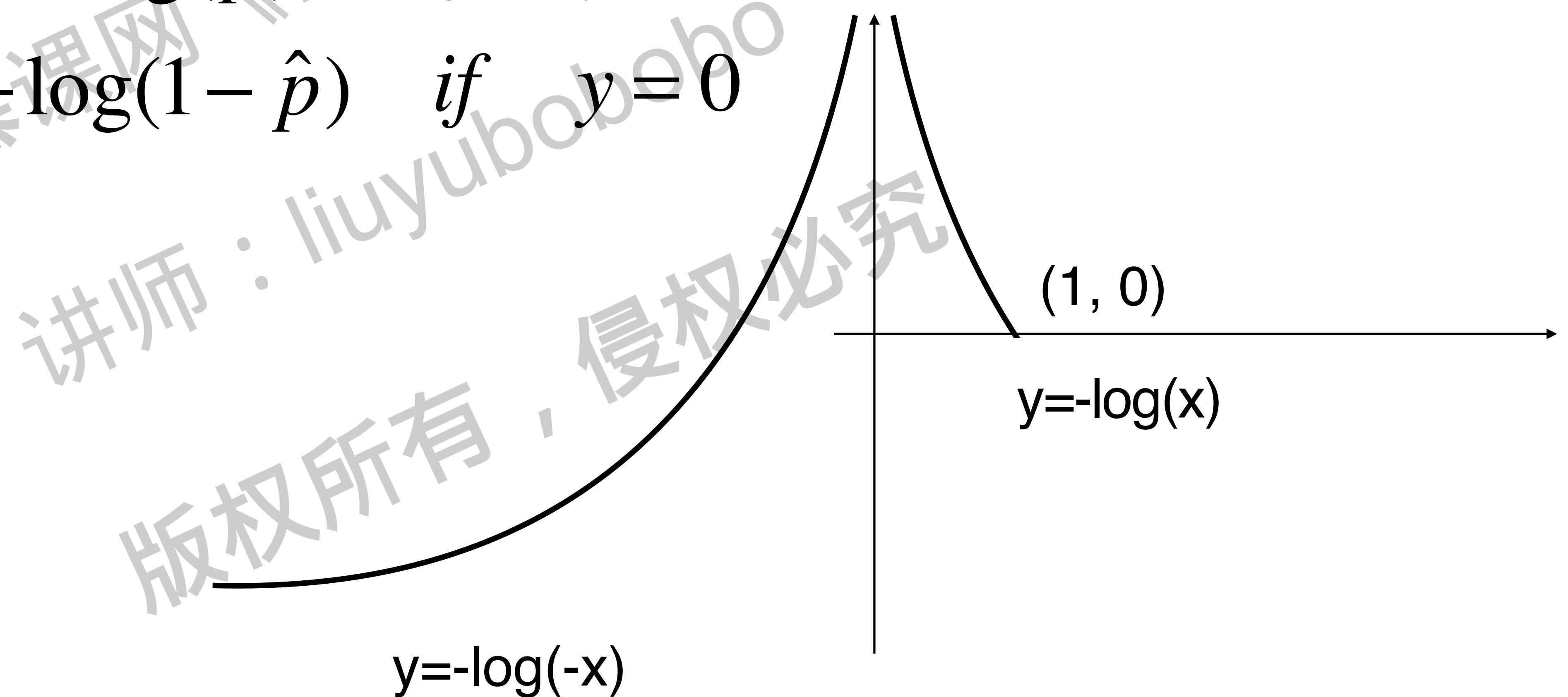
$$\text{cost} = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1 - \hat{p}) & \text{if } y = 0 \end{cases}$$



$y = -\log(x)$

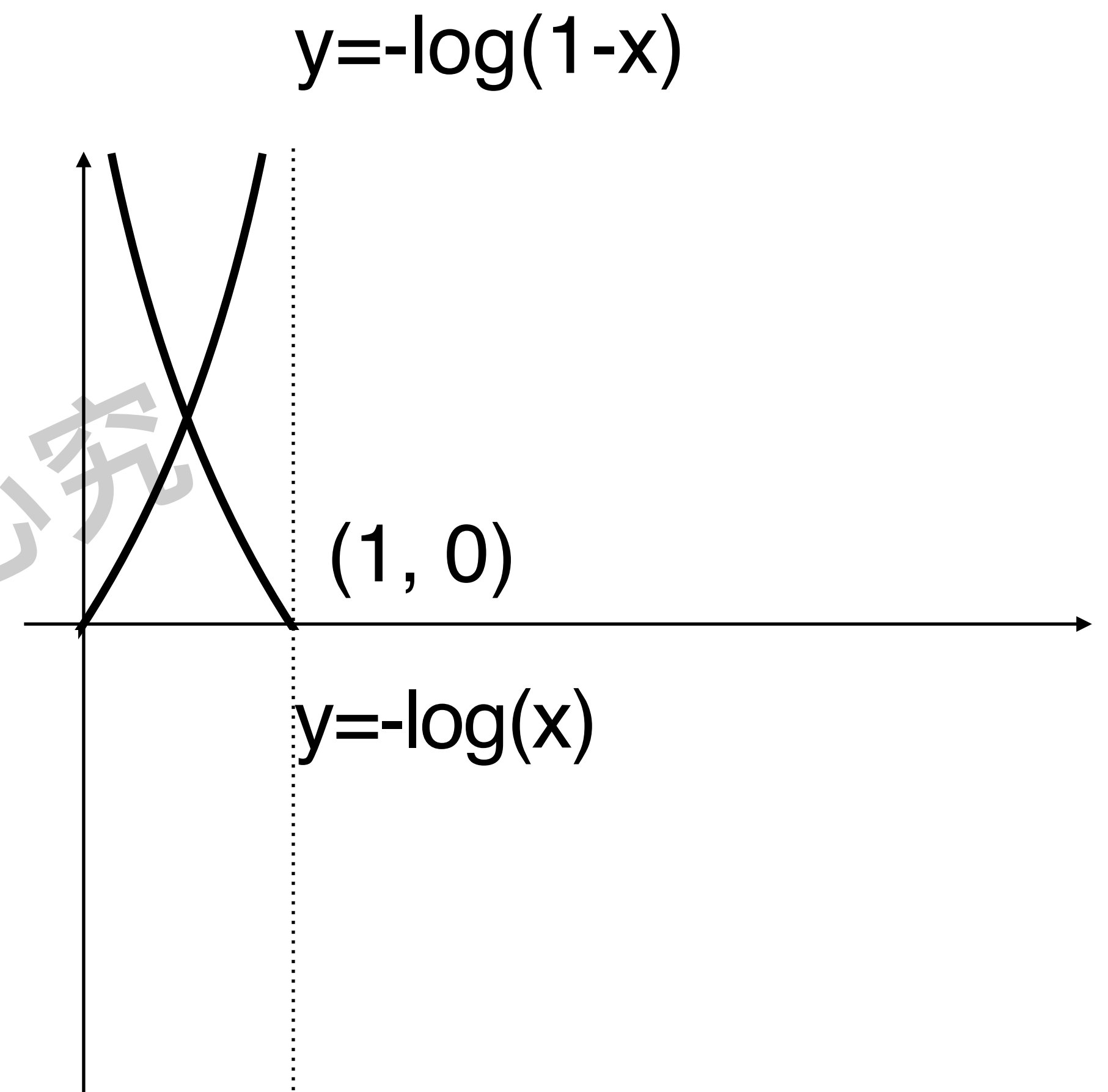
逻辑回归 Logistic Regression

$$\text{cost} = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1 - \hat{p}) & \text{if } y = 0 \end{cases}$$



逻辑回归 Logistic Regression

$$\text{cost} = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1 - \hat{p}) & \text{if } y = 0 \end{cases}$$



逻辑回归 Logistic Regression

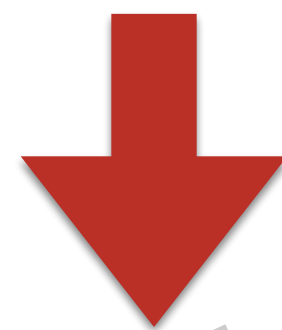
$$\text{cost} = \begin{cases} -\log(\hat{p}) & \text{if } y = 1 \\ -\log(1 - \hat{p}) & \text{if } y = 0 \end{cases}$$



$$\text{cost} = -y \log(\hat{p}) - (1 - y) \log(1 - \hat{p})$$

逻辑回归 Logistic Regression

$$\text{cost} = -y \log(\hat{p}) - (1 - y) \log(1 - \hat{p})$$



$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\hat{p}^{(i)}) + (1 - y^{(i)}) \log(1 - \hat{p}^{(i)})$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\hat{p}^{(i)}) + (1 - y^{(i)}) \log(1 - \hat{p}^{(i)})$$

$$\hat{p}^{(i)} = \sigma(X_b^{(i)} \theta) = \frac{1}{1 + e^{-X_b^{(i)} \theta}}$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\hat{p}^{(i)}) + (1 - y^{(i)}) \log(1 - \hat{p}^{(i)})$$

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\hat{p}^{(i)}) + (1 - y^{(i)}) \log(1 - \hat{p}^{(i)})$$

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

没有公式解，只能使用梯度下降法求解

慕课网《Python3机器学习》

逻辑回归损失函数的梯度

讲师：liuyubobobo

版权所有，侵权必究

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$\nabla J(\theta) = \begin{pmatrix} \frac{\partial J(\theta)}{\partial \theta_0} \\ \frac{\partial J(\theta)}{\partial \theta_1} \\ \dots \\ \frac{\partial J(\theta)}{\partial \theta_n} \end{pmatrix}$$

逻辑回归 Logistic Regression

$$\sigma(t) = \frac{1}{1 + e^{-t}} = (1 + e^{-t})^{-1}$$

$$\sigma(t)' = -(1 + e^{-t})^{-2} \cdot e^{-t} \cdot (-1) = (1 + e^{-t})^{-2} \cdot e^{-t}$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$\sigma(t) = \frac{1}{1 + e^{-t}} = (1 + e^{-t})^{-1} \quad \sigma(t)' = (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$(\log \sigma(t))' = \frac{1}{\sigma(t)} \cdot \sigma(t)' = \frac{1}{\sigma(t)} \cdot (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$= \frac{1}{(1 + e^{-t})^{-1}} \cdot (1 + e^{-t})^{-2} \cdot e^{-t} = (1 + e^{-t})^{-1} \cdot e^{-t}$$

逻辑回归 Logistic Regression

$$\sigma(t) = \frac{1}{1 + e^{-t}} = (1 + e^{-t})^{-1}$$

$$(\log \sigma(t))' = (1 + e^{-t})^{-1} \cdot e^{-t}$$

$$= \frac{e^{-t}}{1 + e^{-t}} = \frac{1 + e^{-t} - 1}{1 + e^{-t}} = 1 - \frac{1}{1 + e^{-t}}$$

$$= 1 - \sigma(t)$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$(\log \sigma(t))' = 1 - \sigma(t)$$

$$\frac{d(y^{(i)} \log \sigma(X_b^{(i)} \theta))}{d\theta_j} = y^{(i)} (1 - \sigma(X_b^{(i)} \theta)) \cdot X_j^{(i)}$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$\sigma(t) = \frac{1}{1 + e^{-t}} = (1 + e^{-t})^{-1} \quad \sigma(t)' = (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$(\log(1 - \sigma(t)))' = \frac{1}{1 - \sigma(t)} \cdot (-1) \cdot \sigma(t)' = -\frac{1}{1 - \sigma(t)} \cdot (1 + e^{-t})^{-2} \cdot e^{-t}$$

逻辑回归 Logistic Regression

$$\sigma(t) = \frac{1}{1 + e^{-t}} = (1 + e^{-t})^{-1} \quad \sigma(t)' = (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$(\log(1 - \sigma(t)))' = \frac{1}{1 - \sigma(t)} \cdot (-1) \cdot \sigma(t)' = -\frac{1}{1 - \sigma(t)} \cdot (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$-\frac{1}{1 - \sigma(t)} = -\frac{1}{\frac{1 + e^{-t}}{1 + e^{-t}} - \frac{1}{1 + e^{-t}}} = -\frac{1 + e^{-t}}{e^{-t}}$$

逻辑回归 Logistic Regression

$$\sigma(t) = \frac{1}{1 + e^{-t}} = (1 + e^{-t})^{-1} \quad \sigma(t)' = (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$(\log(1 - \sigma(t)))' = \frac{1}{1 - \sigma(t)} \cdot (-1) \cdot \sigma(t)' = -\frac{1}{1 - \sigma(t)} \cdot (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$-\frac{1}{1 - \sigma(t)} = -\frac{1 + e^{-t}}{e^{-t}} = -\frac{1 + e^{-t}}{e^{-t}} \cdot (1 + e^{-t})^{-2} \cdot e^{-t}$$

$$= -(1 + e^{-t})^{-1} = -\sigma(t)$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$(\log(1 - \sigma(t)))' = -\sigma(t)$$

$$\frac{d((1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta)))}{d\theta_j} = (1 - y^{(i)}) \cdot (-\sigma(X_b^{(i)} \theta)) \cdot X_j^{(i)}$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$\frac{d(y^{(i)} \log \sigma(X_b^{(i)} \theta))}{d\theta_j} = y^{(i)} (1 - \sigma(X_b^{(i)} \theta)) \cdot X_j^{(i)}$$

$$+ \frac{d((1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta)))}{d\theta_j} = (1 - y^{(i)}) \cdot (-\sigma(X_b^{(i)} \theta)) \cdot X_j^{(i)}$$

$$y^{(i)} X_j^{(i)} - y^{(i)} \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)} - \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)} + y^{(i)} \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)}$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$y^{(i)} X_j^{(i)} - y^{(i)} \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)} - \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)} + y^{(i)} \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)}$$

$$= y^{(i)} X_j^{(i)} - \sigma(X_b^{(i)} \theta) \cdot X_j^{(i)} = (y^{(i)} - \sigma(X_b^{(i)} \theta)) \cdot X_j^{(i)}$$

$$\frac{J(\theta)}{\theta_j} = \frac{1}{m} \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) X_j^{(i)}$$

逻辑回归 Logistic Regression

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(\sigma(X_b^{(i)} \theta)) + (1 - y^{(i)}) \log(1 - \sigma(X_b^{(i)} \theta))$$

$$\frac{J(\theta)}{\theta_j} = \frac{1}{m} \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) X_j^{(i)}$$

$$= \frac{1}{m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) X_j^{(i)}$$

逻辑回归 Logistic Regression

$$\frac{J(\theta)}{\theta_j} = \frac{1}{m} \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) X_j^{(i)}$$

$$\nabla J(\theta) = \begin{pmatrix} \frac{\partial J}{\partial \theta_0} \\ \frac{\partial J}{\partial \theta_1} \\ \frac{\partial J}{\partial \theta_2} \\ \dots \\ \frac{\partial J}{\partial \theta_n} \end{pmatrix} = \frac{1}{m} \cdot \begin{pmatrix} \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) \\ \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) \cdot X_1^{(i)} \\ \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) \cdot X_2^{(i)} \\ \dots \\ \sum_{i=1}^m (\sigma(X_b^{(i)} \theta) - y^{(i)}) \cdot X_n^{(i)} \end{pmatrix} = \frac{1}{m} \cdot \begin{pmatrix} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_1^{(i)} \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_2^{(i)} \\ \dots \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_n^{(i)} \end{pmatrix}$$

逻辑回归 Logistic Regression

$$\nabla J(\theta) = \frac{1}{m} \cdot \begin{pmatrix} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_1^{(i)} \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_2^{(i)} \\ \dots \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_n^{(i)} \end{pmatrix}$$

回忆线性回归

$$\nabla J(\theta) = \frac{2}{m} \cdot \begin{pmatrix} \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \\ \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \cdot X_1^{(i)} \\ \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \cdot X_2^{(i)} \\ \dots \\ \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \cdot X_n^{(i)} \end{pmatrix}$$

逻辑回归 Logistic Regression

回忆线性回归

$$\nabla J(\theta) = \frac{2}{m} \cdot \begin{pmatrix} \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \\ \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \cdot X_1^{(i)} \\ \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \cdot X_2^{(i)} \\ \dots \\ \sum_{i=1}^m (X_b^{(i)} \theta - y^{(i)}) \cdot X_n^{(i)} \end{pmatrix} = \frac{2}{m} \cdot X_b^T \cdot (X_b \theta - y)$$

逻辑回归 Logistic Regression

$$\nabla J(\theta) = \frac{1}{m} \cdot \begin{pmatrix} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_1^{(i)} \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_2^{(i)} \\ \dots \\ \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) \cdot X_n^{(i)} \end{pmatrix} = \frac{1}{m} \cdot X_b^T \cdot (\sigma(X_b \theta) - y)$$

慕课网《Python3机器学习》

实现我们自己的逻辑回归

讲师：liuyubon0199

版权所有，侵权必究

实践：实现我们自己的逻辑回归

讲师：liuyuboboo
版权所有，侵权必究

慕课网《Python3机器学习》

决策边界

讲师：liuyuboboo

版权所有，侵权必究

逻辑回归 Logistic Regression

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} < 0.5 \end{cases}$$

逻辑回归 Logistic Regression

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 \\ 0, & \hat{p} < 0.5 \end{cases}$$

$t > 0$ 时, $p > 0.5$

$t < 0$ 时, $p < 0.5$

逻辑回归 Logistic Regression

$$\hat{p} = \sigma(\theta^T \cdot x_b) = \frac{1}{1 + e^{-\theta^T \cdot x_b}}$$

$$\hat{y} = \begin{cases} 1, & \hat{p} \geq 0.5 & \theta^T \cdot x_b \geq 0 \\ 0, & \hat{p} < 0.5 & \theta^T \cdot x_b < 0 \end{cases}$$

决策边界

$$\theta^T \cdot x_b = 0$$

逻辑回归 Logistic Regression

决策边界

如果X有两个特征

$$\theta^T \cdot x_b = 0$$

$$\theta_0 + \theta_1 x_1 + \theta_2 x_2 = 0$$

$$x_2 = \frac{-\theta_0 - \theta_1 x_1}{\theta_2}$$

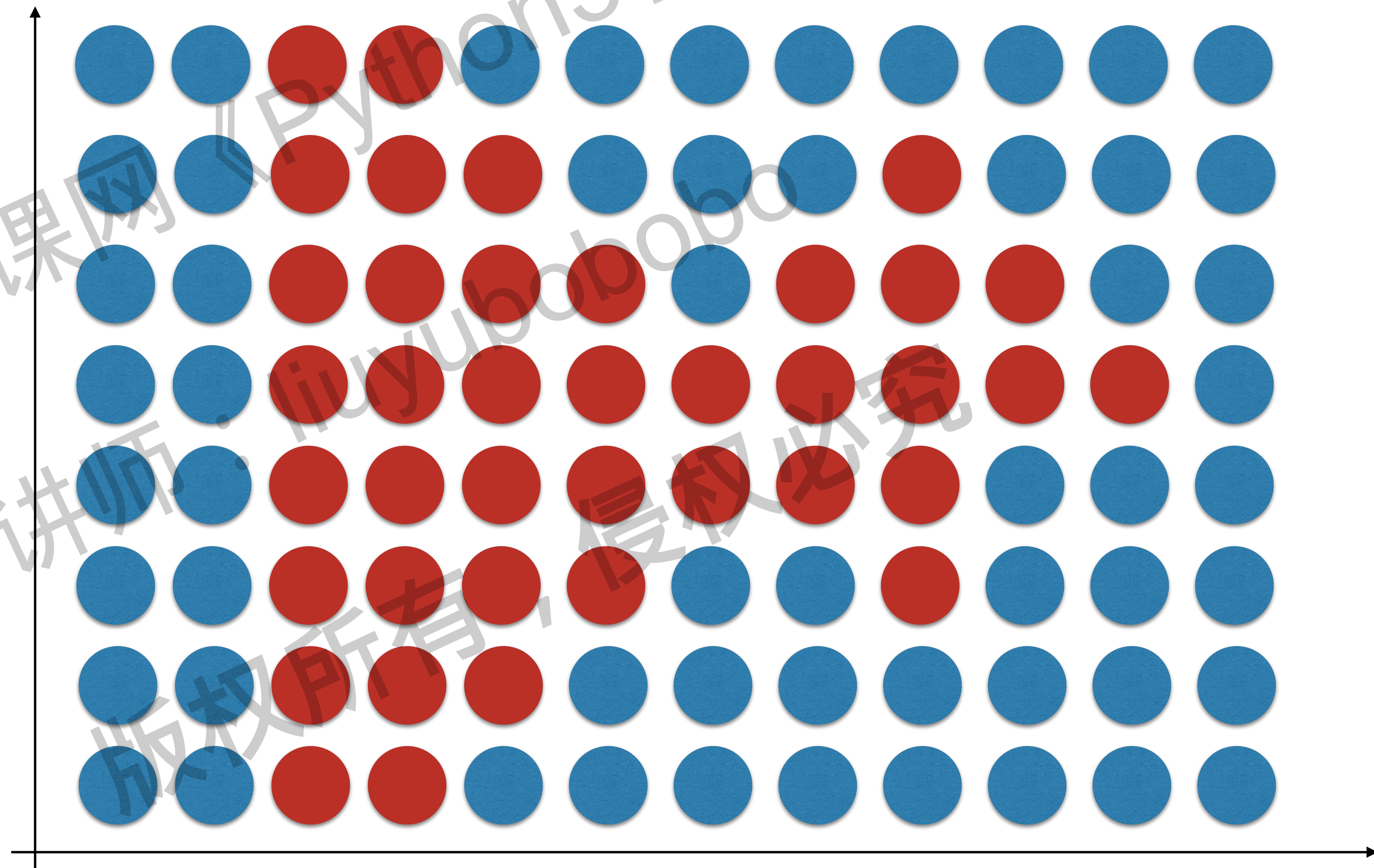
慕课网《Python3机器学习》

实践：绘制决策边界

讲师：liuyuboboo

版权所有，侵权必究

不规则的决策边界的绘制方法



实践：另一种绘制决策边界的方法

讲师：liuyuboboo
版权所有，侵权必究

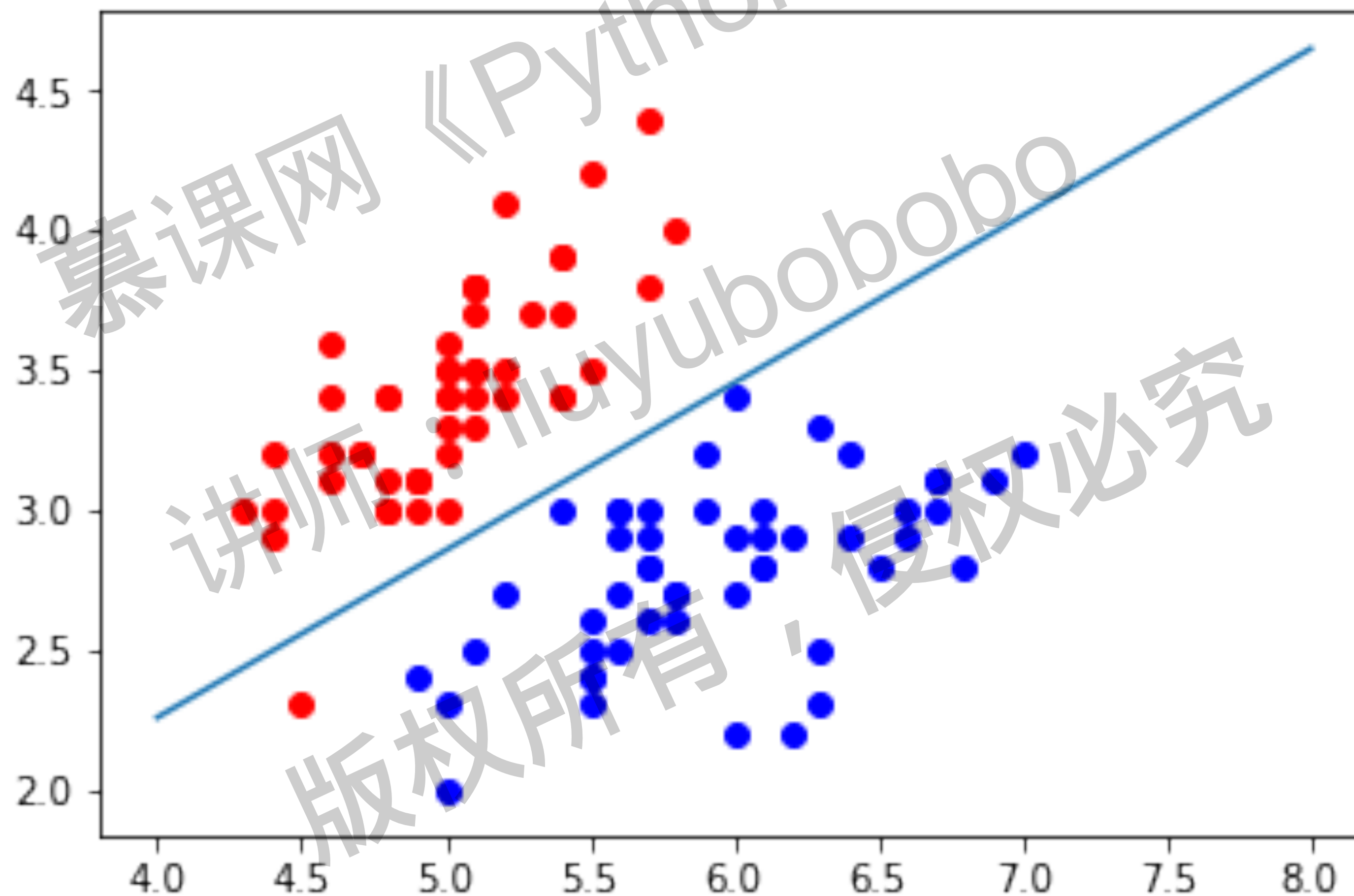
慕课网《Python3机器学习》

逻辑回归中使用多项式项

讲师：liuyubobobo

版权所有，侵权必究

逻辑回归 Logistic Regression

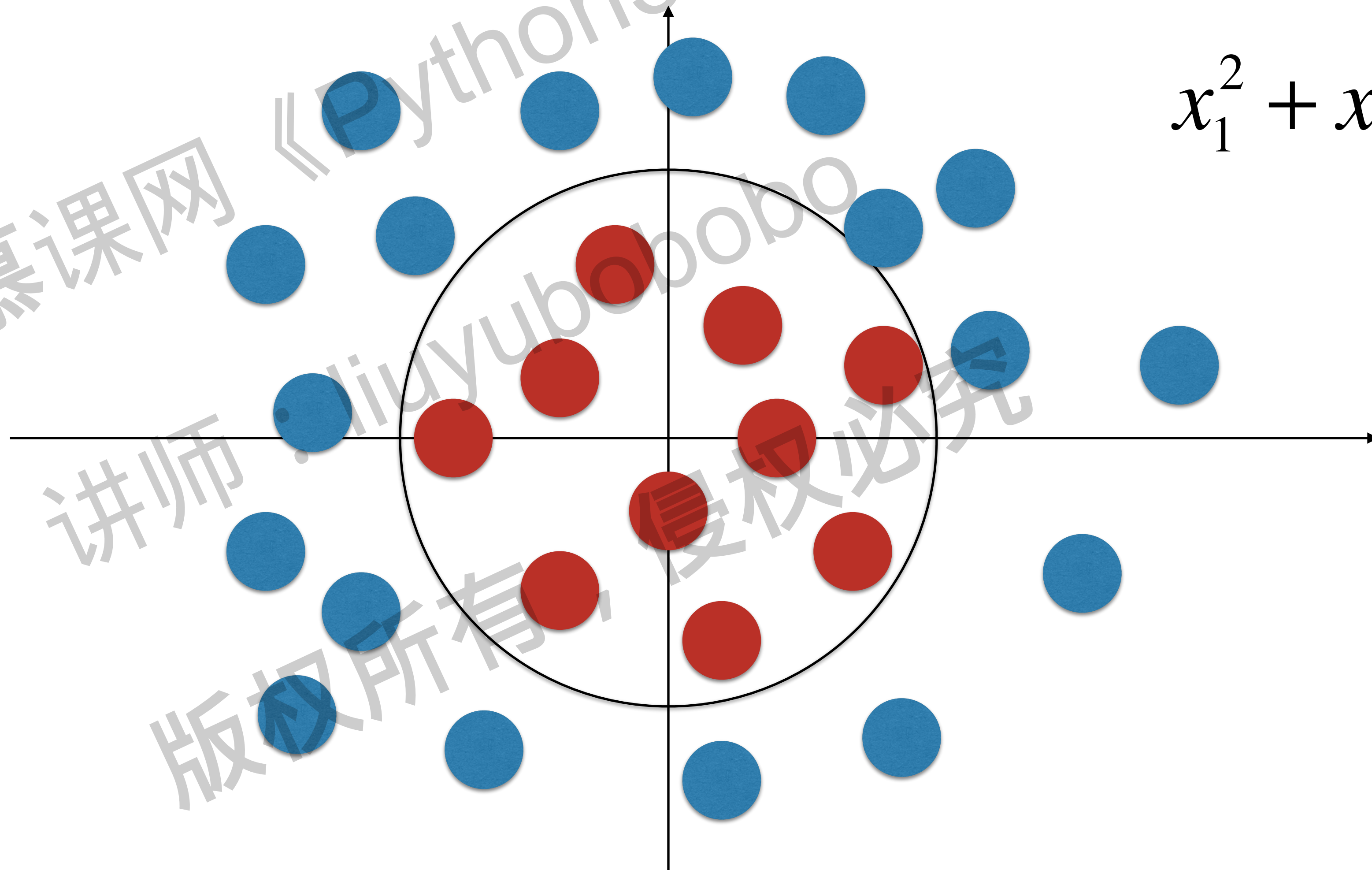


决策边界

$$\theta^T \cdot x_b = 0$$

逻辑回归 Logistic Regression

$$x_1^2 + x_2^2 - r^2 = 0$$



实践：在逻辑回归中添加多项式项

讲师：liuyuboboo
版权所有，侵权必究

慕课网《Python3机器学习》

scikit-learn中的逻辑回归

讲师：liuyubobobo

版权所有，侵权必究

逻辑回归中使用正则化

$$J(\theta) + \alpha L_2$$

$$J(\theta) + \alpha L_1$$

逻辑回归中使用正则化

$$J(\theta) + \alpha L_2$$

$$C \cdot J(\theta) + L_1$$

$$J(\theta) + \alpha L_1$$

$$C \cdot J(\theta) + L_2$$

scikit-learn中使用的方式

实践：scikit-learn中的逻辑回归

讲师：liuyubobobo

版权所有，侵权必究

慕课网《Python3机器学习》

OvR 和 OvO

讲师：liuyubobobo

版权所有，侵权必究

逻辑回归 Logistic Regression

逻辑回归只可以解决二分类问题

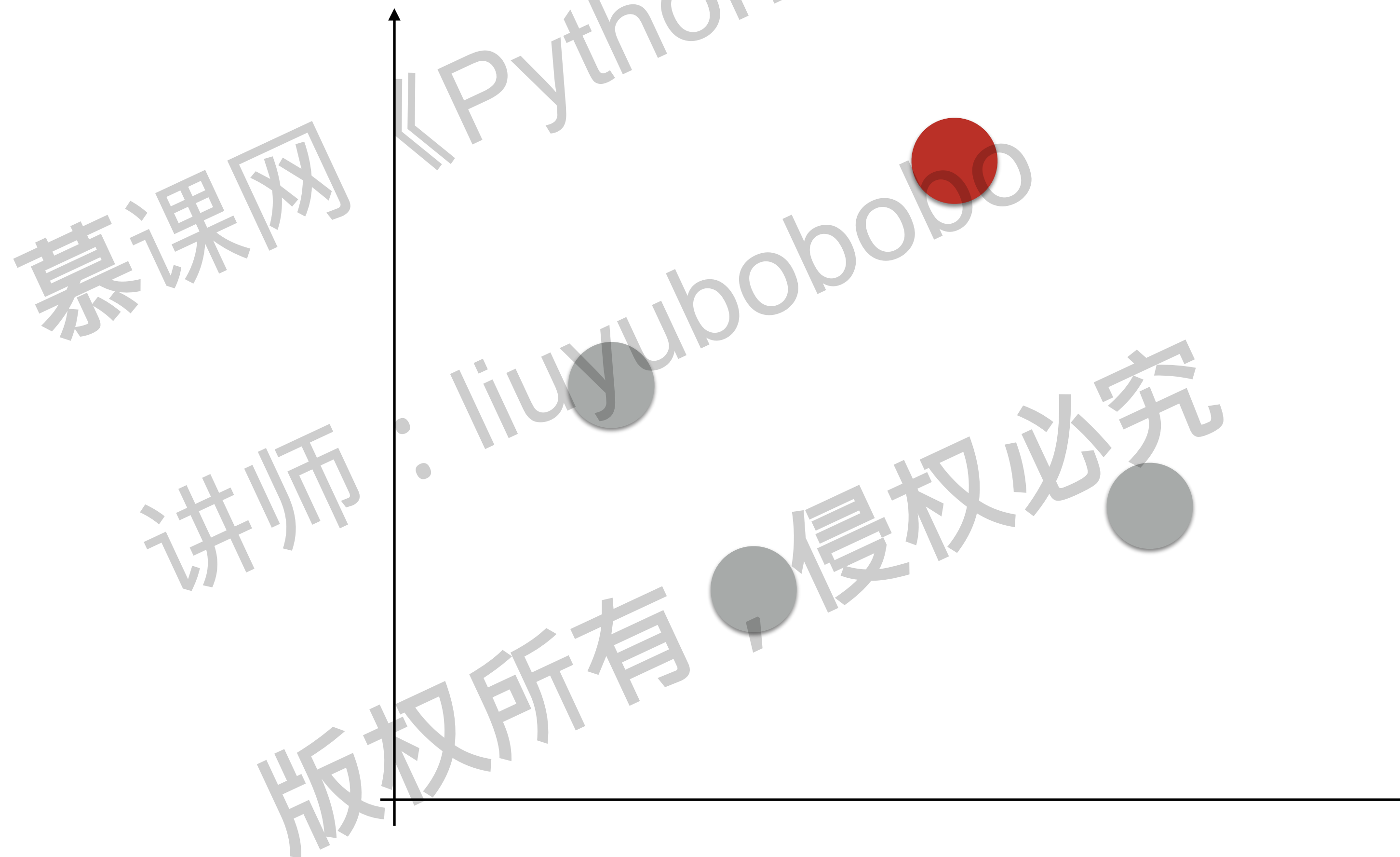
解决多分类问题：

- OvR
- OvO

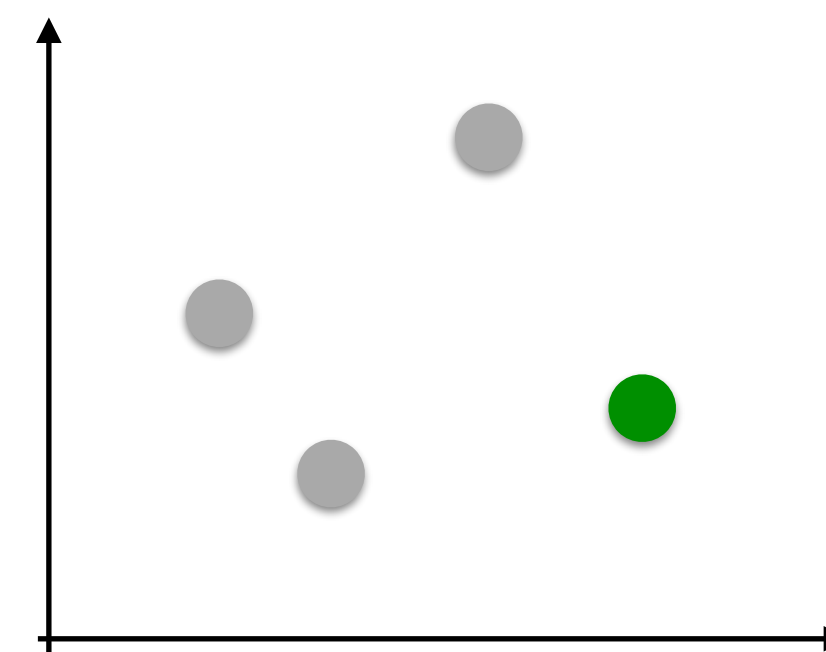
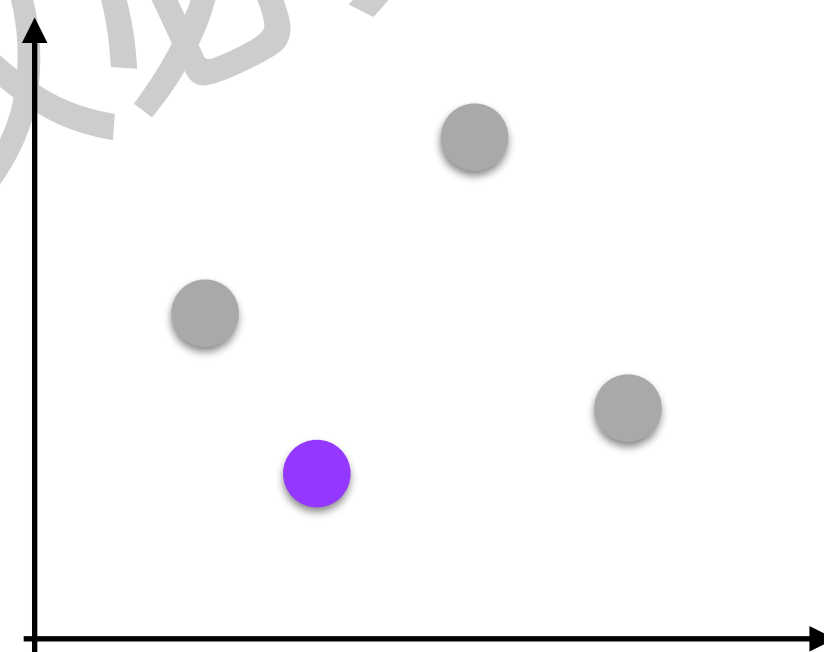
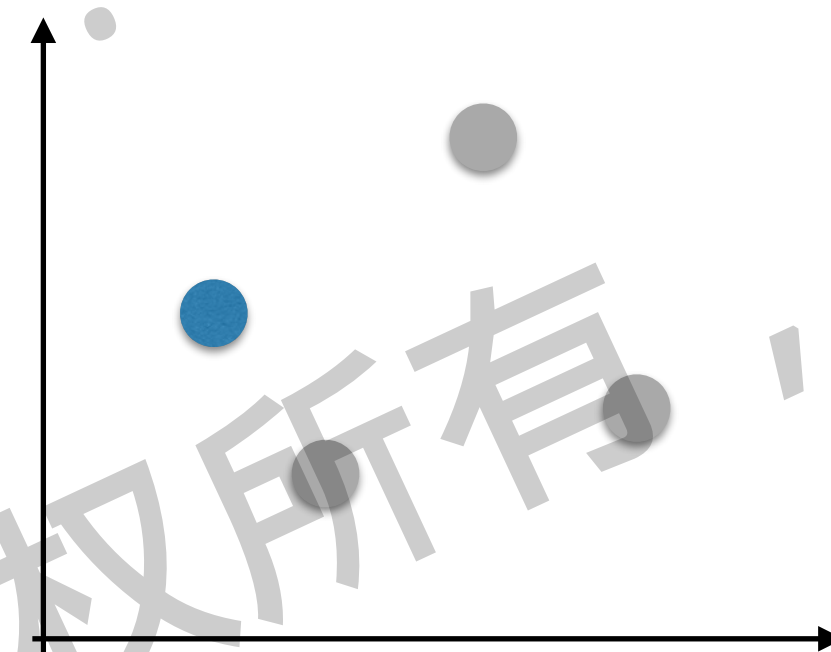
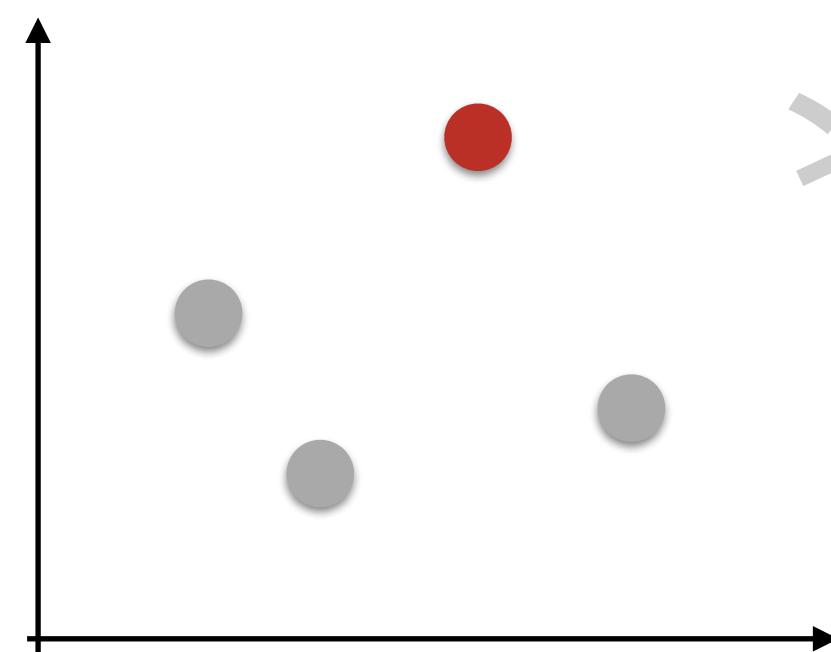
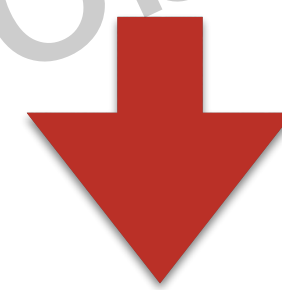
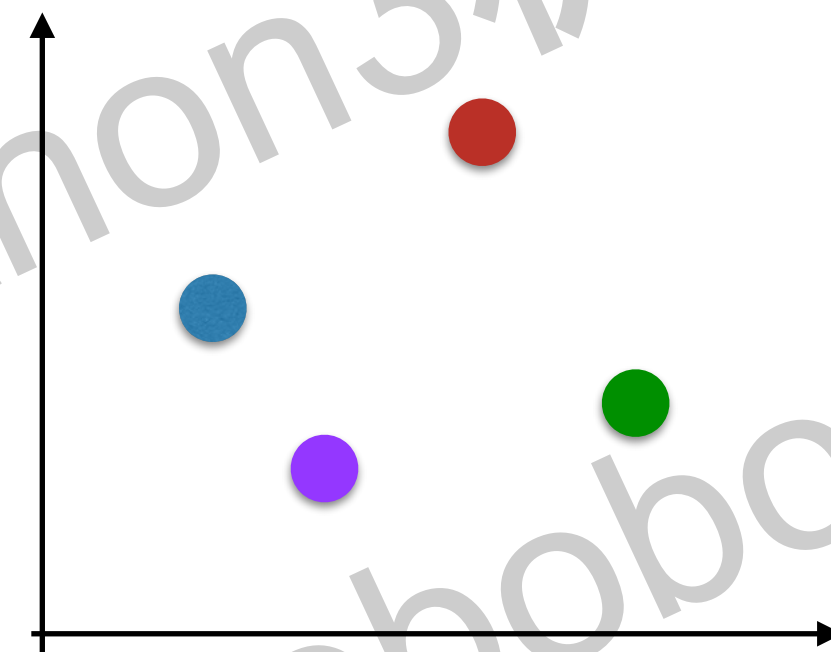
OvR (One vs Rest)



OvR (One vs Rest)



OvR (One vs Rest)



n个类别就进行n次分类，选择分类得分最高的

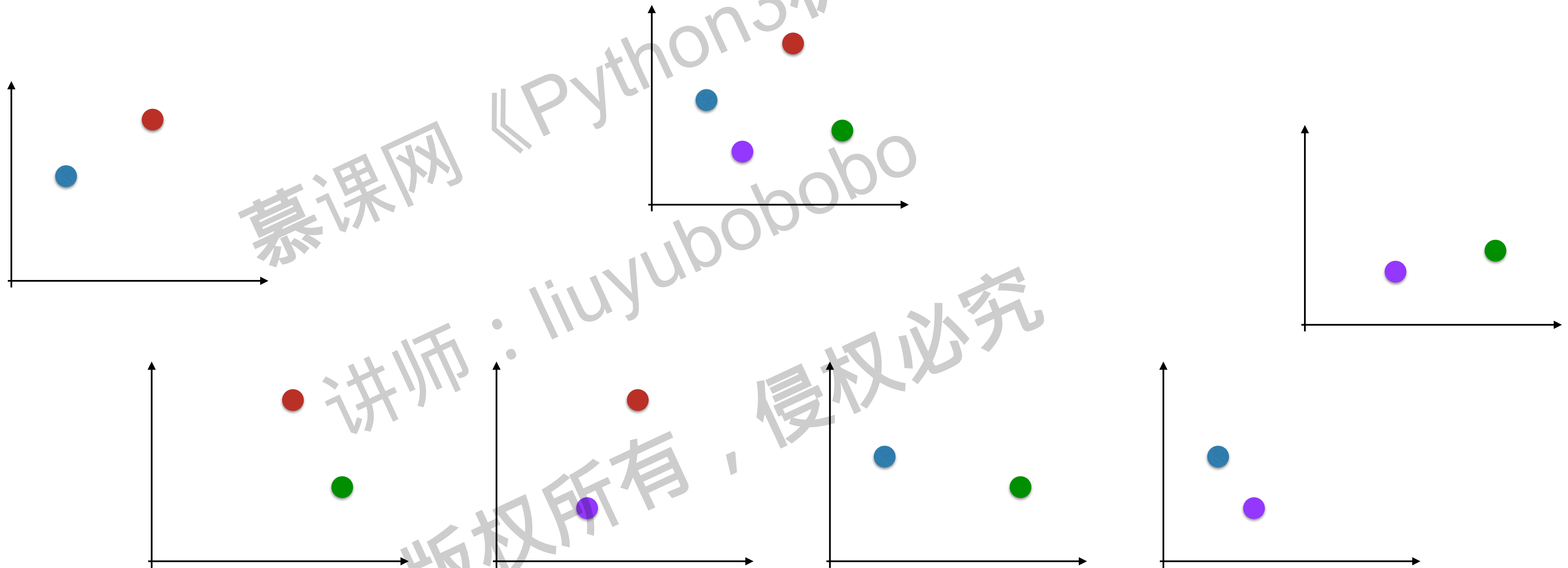
OvO (One vs One)



OvO (One vs One)



OvO (One vs One)



n 个类别就进行 $C(n,2)$ 次分类，选择赢数最高的分类

慕课网《Python3机器学习》

实践：OvR and OvO

讲师：liuyubobobo

版权所有，侵权必究

其他

欢迎大家关注我的个人公众号：是不是很酷



Python 3 玩儿转机器学习

讲师：liuyubobobo

版权所有 侵权必究
liuyubobobo