

如何解决多元分类问题

多元逻辑回归、OvR、OvO

小胖

目录

ONE 多元逻辑回归

隐含变量模型

TWO 降维到二元分类问题

OvR、OvO

THREE 代码实现

scikit-learn

多元逻辑回归

多元分类问题

二元分类问题

在生活中，我们经常会遇到二元选择问题



在数学上，如果用 y 表示选择的结果，那么 y 只有两个可能的取值：0或者1

多元分类问题

在生活中，我们还会遇到多元选择问题



在数学上，如果用 y 表示选择的结果，那么 y 有三个可能的取值：0、1、2

多元逻辑回归

数学细节

从隐含变量模型出发，
推导出各个类别的概率
(模型预测的概率)

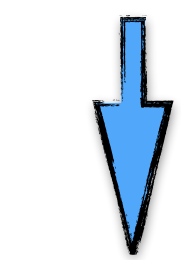
定义k个隐含变量模型

个体属于某个类别，当且仅
当这个类别对它的效用最大

推导出各个类别概率分布

$$\begin{cases} Y_{i,0}^* = X_i \theta_0 + \varepsilon_0 \\ Y_{i,1}^* = X_i \theta_1 + \varepsilon_1 \\ \vdots \\ Y_{i,k-1}^* = X_i \theta_{k-1} + \varepsilon_{k-1} \end{cases}$$

其中 $Y_{m,l}^*$ 表示类别l对个体m的效用
 ε_i 是随机扰动项，服从标准的
类型1极端值分布



$$\begin{cases} P(Y_i = 0) = P(Y_{i,0}^* = \max_j Y_{i,j}^*) \\ P(Y_i = 1) = P(Y_{i,1}^* = \max_j Y_{i,j}^*) \\ \vdots \\ P(Y_i = k-1) = P(Y_{i,k-1}^* = \max_j Y_{i,j}^*) \end{cases}$$



$$\begin{cases} P(Y_i = 1) = P(Y_i = 0)e^{X_i \beta_1} = \frac{e^{X_i \beta_1}}{(1 + \sum_{j=1}^{k-1} e^{X_i \beta_j})} \\ P(Y_i = 2) = P(Y_i = 0)e^{X_i \beta_2} = \frac{e^{X_i \beta_2}}{(1 + \sum_{j=1}^{k-1} e^{X_i \beta_j})} \\ \dots \\ P(Y_i = 0) = \frac{1}{(1 + \sum_{j=1}^{k-1} e^{X_i \beta_j})} \end{cases}$$

多元逻辑回归

数学细节

模型参数的估计公式

各类别概率分布

$$\begin{cases} P(Y_i = 1) = P(Y_i = 0)e^{X_i\beta_1} = \frac{e^{X_i\beta_1}}{(1 + \sum_{j=1}^{k-1} e^{X_i\beta_j})} \\ P(Y_i = 2) = P(Y_i = 0)e^{X_i\beta_2} = \frac{e^{X_i\beta_2}}{(1 + \sum_{j=1}^{k-1} e^{X_i\beta_j})} \\ \dots \\ P(Y_i = 0) = \frac{1}{(1 + \sum_{j=1}^{k-1} e^{X_i\beta_j})} \end{cases}$$



定义模型似然函数

$$\begin{aligned} L &= P(Y | \beta) = \prod_i \prod_{j=0}^{k-1} P(Y_i = j)^{1_{\{Y_i=j\}}} \\ \ln L &= \sum_i \sum_{j=0}^{k-1} 1_{\{Y_i=j\}} \ln P(Y_i = j) \end{aligned}$$



利用最大似然估计法估计参数

$$\hat{\beta} = \operatorname{argmax}_{\beta} L = \operatorname{argmax}_{\beta} \ln L = \operatorname{argmax}_{\beta} \ln P(Y | \beta)$$

目录

ONE 多元逻辑回归

隐含变量模型

TWO 降维到二元分类问题

OvR、OvO

THREE 代码实现

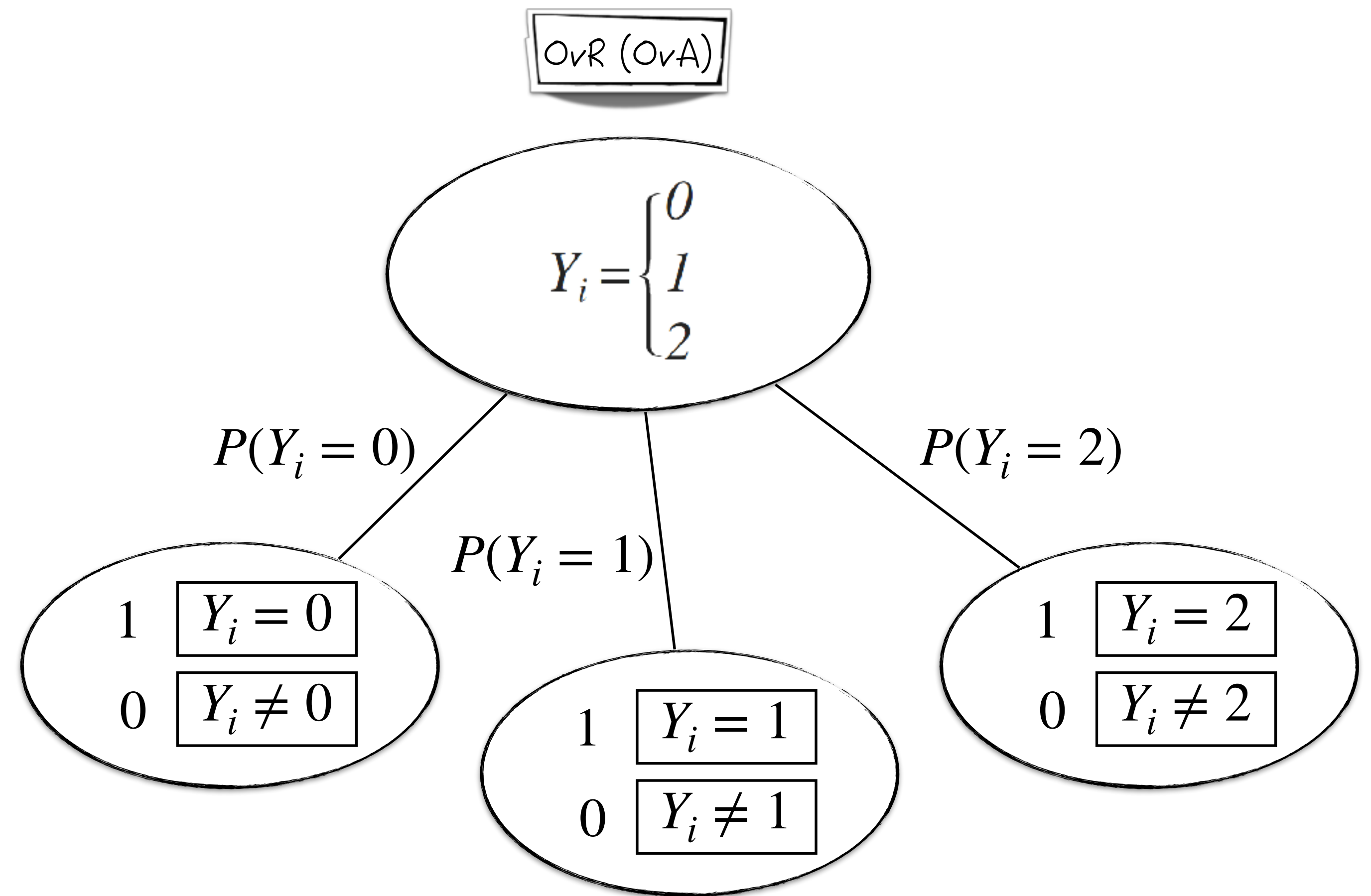
scikit-learn

降维到二元分类问题

OvR

可以将多元分类问题分解为多个二元分类问题来解决，解决的策略通常有两种

- OvR (OvA) : One-vs.-rest
- OvO: One-vs.-one



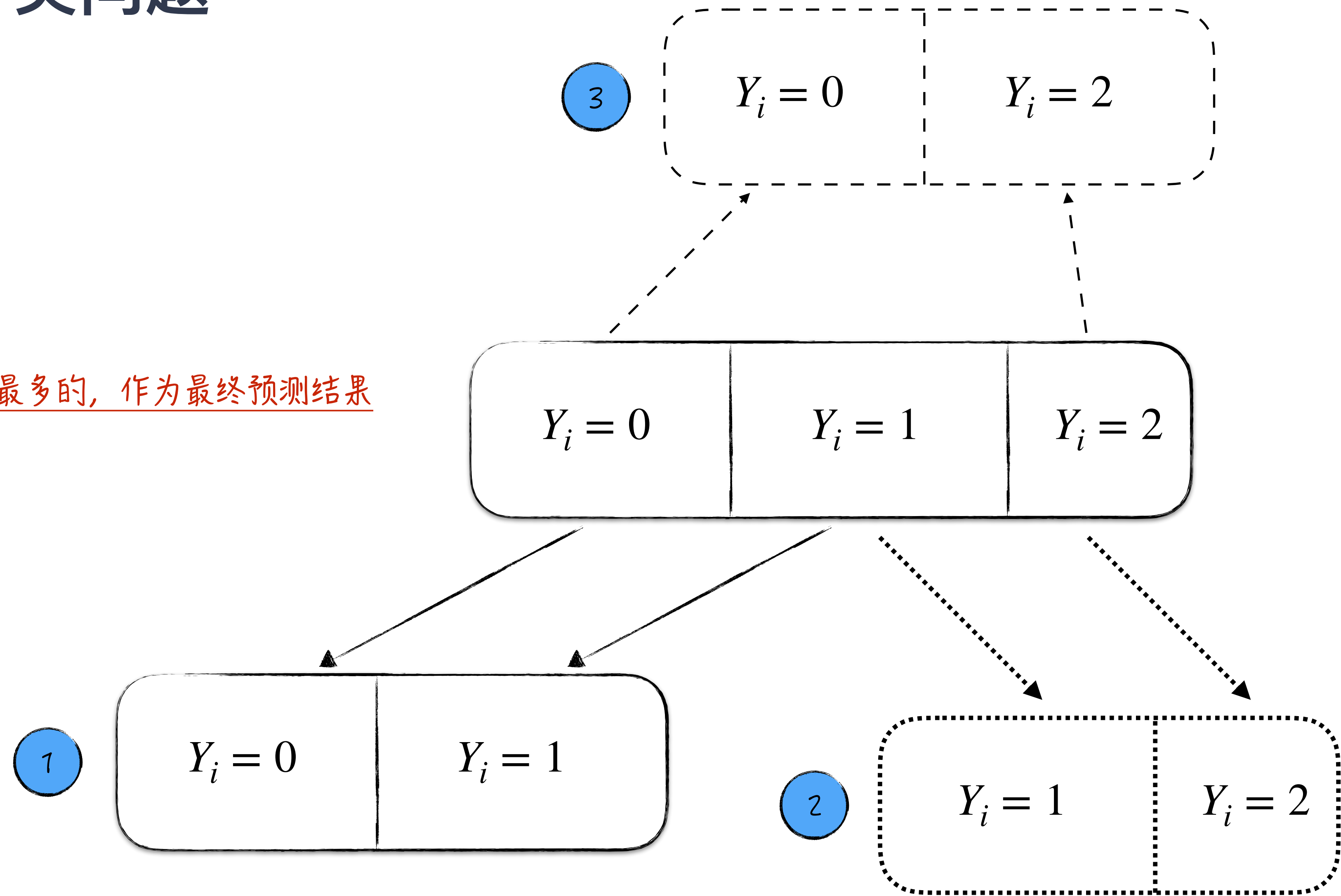
选择概率值最大的，作为最终预测结果

比如 $P(Y_i = 0) > P(Y_i = 1) > P(Y_i = 2) \Rightarrow \hat{Y}_i = 0$

降维到二元分类问题

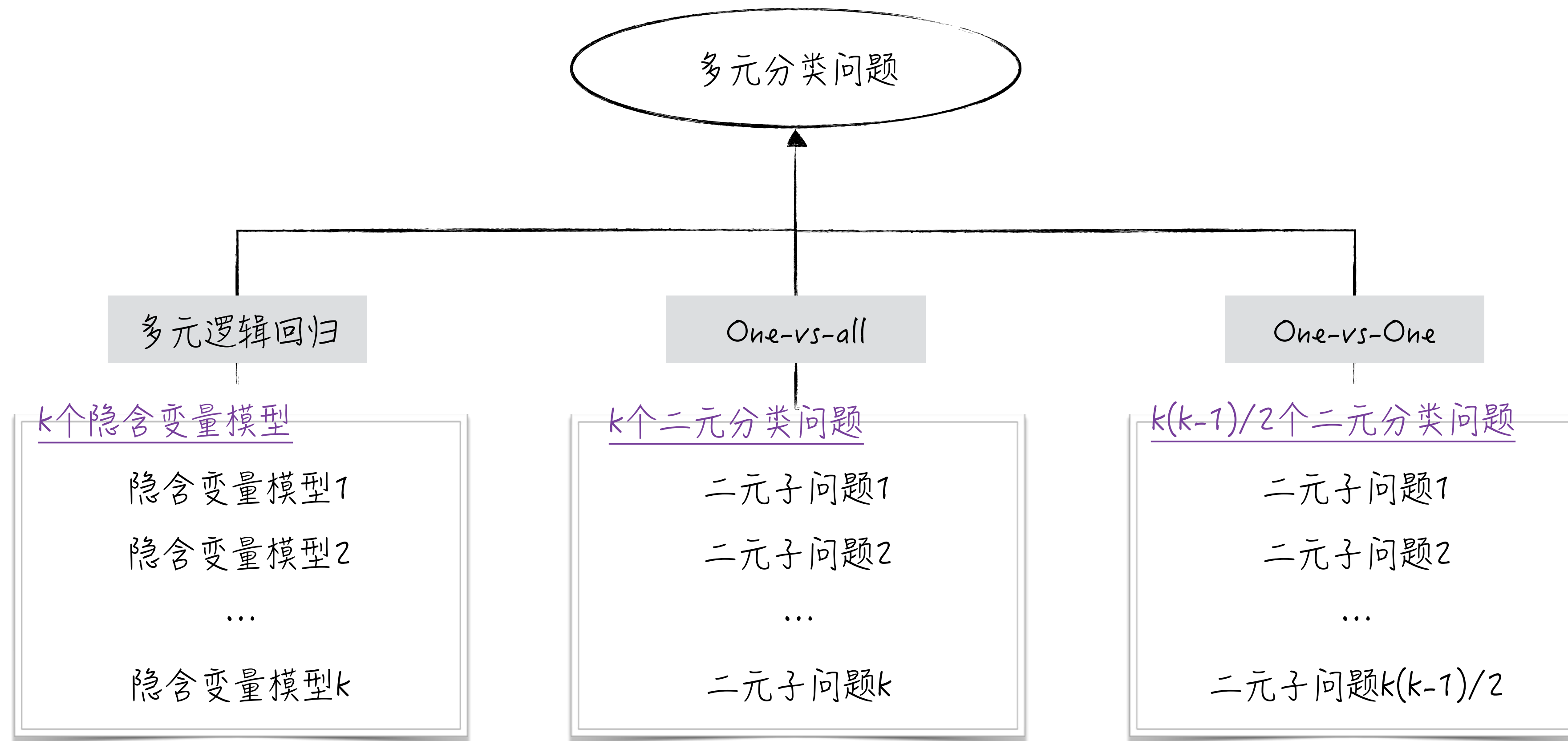
OvO

“获胜”次数最多的，作为最终预测结果



降维到二元分类问题

OvO



目录

ONE 多元逻辑回归

隐含变量模型

TWO 降维到二元分类问题

OvR、OvO

THREE 代码实现

scikit-learn

THANK YOU

—