所用的模型：对数几率回归

$$\ln\frac{p(y=1\,|\,\mathbf{x})}{p(y=0\,|\,\mathbf{x})}=\boldsymbol{\omega}^T\mathbf{x}+b$$

$$p(y=1\,|\,\mathbf{x})=\frac{e^{\boldsymbol{\omega}^T\mathbf{x}+b}}{1+e^{\boldsymbol{\omega}^T\mathbf{x}+b}}$$

$$p(y=0\,|\,\mathbf{x})=\frac{1}{1+e^{\boldsymbol{\omega}^T\mathbf{x}+b}}$$

对于输入的 $\mathbf{x}$，计算 $\frac{p(y=1\,|\,\mathbf{x})}{p(y=0\,|\,\mathbf{x})}$ 与设定的阈值比较，大于阈值的为正类，反之为反类

参数确定：

为了便于表示和计算，令 $\boldsymbol{\beta}=(\boldsymbol{\omega};\ b),\hat{\mathbf{x}}=(\mathbf{x};1)$， $p_i(\boldsymbol{\beta})=p(y=i\,|\,\hat{\mathbf{x}};\boldsymbol{\beta})$， m 为样本数，利用最大化对数似然估计：

$$l(\boldsymbol{\beta})=\sum_{i=1}^{m}\ln p(y_i\,||\,\hat{\mathbf{x}}_{\mathbf{i}};\boldsymbol{\beta})$$

等价于最小化：

$$l(\boldsymbol{\beta})=\sum_{i=1}^{m}(-y_i\boldsymbol{\beta}^T\hat{\mathbf{x}}_{\mathbf{i}}+\ln(1+e^{\boldsymbol{\beta}^T\hat{\mathbf{x}}_{\mathbf{i}}}))$$

故所要求的参数向量为：

$$\boldsymbol{\beta}^*=\arg\min_{\boldsymbol{\beta}} l(\boldsymbol{\beta})$$

利用梯度下降算法得到最优解：

（1）取初值 $\boldsymbol{\beta}^{(0)}$=[0.1;0.1;0.1;0.1;0.1]， 置 k=0；

（2）计算$l(\boldsymbol{\beta}^{(k)})=\sum_{i=1}^{m}(-y_i\boldsymbol{\beta}^{(k)T}\hat{\mathbf{x}}_{\mathbf{i}}+\ln(1+e^{\boldsymbol{\beta}^{(k)T}\hat{\mathbf{x}}_{\mathbf{i}}}))$；

（3）计算梯度$\frac{\partial l(\boldsymbol{\beta}^{(k)})}{\partial\boldsymbol{\beta}^{(k)}}=-\sum_{i=1}^{m}\hat{\mathbf{x}}_i(y_i-p_1(\hat{\mathbf{x}}_i;\boldsymbol{\beta}^{(k)}))$，并且一维搜索确定$\eta_k$得到$\min_{\eta_k\geq 0} l(\boldsymbol{\beta}^{(k)}-\eta_k\times(\frac{\partial l(\boldsymbol{\beta}^{(k)})}{\partial\boldsymbol{\beta}^{(k)}}))$

（4）$\boldsymbol{\beta}^{(k+1)}=\boldsymbol{\beta}^{(k)}-\eta_k\times(\frac{\partial l(\boldsymbol{\beta}^{(k)})}{\partial\boldsymbol{\beta}^{(k)}})$，并计算$l(\boldsymbol{\beta}^{(k+1)})$，当$\| l(\boldsymbol{\beta}^{(k+1)})-l(\boldsymbol{\beta}^{(k)})\|<\varepsilon$时，令$\boldsymbol{\beta}^{*}=\boldsymbol{\beta}^{(k)}$，停止迭代

（5）否则置 k=k+1,返回（2）