2.2 Softmax分类器

CSDN学院

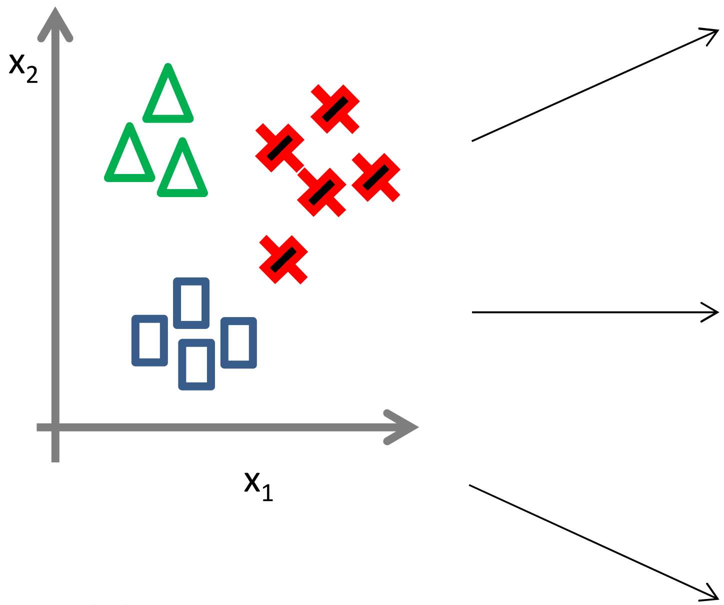
2017年10月

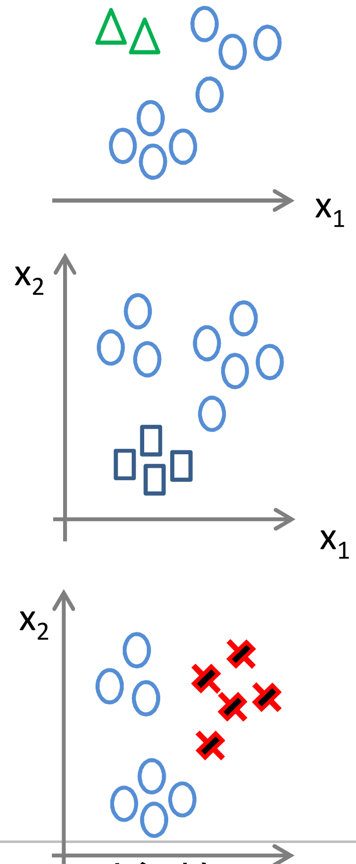
大纲

- Logistic回归基本原理
- 多类Logistic回归
- Scikit learn 中的Logistic回归实现
- 分类模型的评价
- 模型选择与参数调优
- 案例分析

▶ 多类分类任务

一对所有(One-vs-all /One-vs-rest)：





$$f_{\mathbf{w}}^{c}(\mathbf{x}) = p(y = c \mid \mathbf{x}, \mathbf{W}), \quad c = 1, 2, 3$$

如果是正则LR，每类的模型都有自己正则参数

One-vs-all

对每个类别c，训练一个logistic回归分类器$f_{\mathbf{w}}^{c}\left(\mathbf{x}\right)$ ，预测概率 $y=c$.

对新的输入$\mathbf{x}$，选择使得$f_{\mathbf{w}}^{c}\left(\mathbf{x}\right)$最大的类别作为其预测： $\max_{c} f_{\mathbf{w}}^{c}\left(\mathbf{x}\right)$

► Softmax分类器

不止于代码

- 从sigmoid（对应二项分布）扩展为softmax函数（对应多项分布 Cat）：

$$p(y=c\mid\mathbf{x},\mathbf{W})=\mathrm{Cat}\left(y\mid\mathcal{S}\left(\mathbf{W}^T\mathbf{x}\right)\right)$$

- Softmax 函数类似取最大函数： $\mathcal{S}(\boldsymbol{\eta})_c=\dfrac{\exp(\eta_c)}{\sum_{c'=1}^{C}\exp(\eta_{c'})}$

- 综合起来： $\boxed{p(y=c\mid\mathbf{x},\mathbf{W})=\dfrac{\exp(\mathbf{w}_c^T\mathbf{x})}{\sum_{c'=1}^{C}\exp(\mathbf{w}_{c'}^T\mathbf{x})}}$

▶ Softmax回归

$$p(y=c\,|\,\mathbf{x},\mathbf{W})=\frac{\exp(\mathbf{w}_c^T\mathbf{x})}{\sum_{c'=1}^{C}\exp(\mathbf{w}_{c'}^T\mathbf{x})}$$

- 引入记号：$\mu_{ic}=p\left(y_i=c\,|\,\mathbf{x}_i,\mathbf{W}\right)=\mathcal{S}\left(\boldsymbol{\eta}_i\right)_c$

  $\boldsymbol{\eta}_i=\mathbf{W}^T\mathbf{x}_i \qquad C\times 1$ vector

  $y_{ic}=\mathbb{I}\left(y_i=c\right)$

- 则负似然函数为：$J\left(\mathbf{W}\right)=NLL\left(\mathbf{W}\right)=-l\left(\mathbf{W}\right)=-\log\prod_{i=1}^{N}\prod_{c=1}^{C}\mu_{ic}^{y_{ic}}=-\sum_{i=1}^{N}\sum_{c=1}^{C}y_{ic}\log\mu_{ic}$

$$=-\sum_{i=1}^{N}\left[\left(\sum_{c=1}^{C}y_{ic}\mathbf{w}_c^T\mathbf{x}_i\right)-\log\left(\sum_{c'=1}^{C}\exp\left(\mathbf{w}_{c'}^T\mathbf{x}_i\right)\right)\right]$$

  – 梯度：$\mathbf{g}=\left[\nabla J\left(\mathbf{w}_1\right),\ldots,\nabla J\left(\mathbf{w}_C\right)\right]=\left[\mathbf{g}_1,\ldots,\mathbf{g}_C\right]$

  $$\mathbf{g}_c=\sum_{i=1}^{N}\left(\mu_{ic}-y_{ic}\right)\mathbf{x}_i$$

  – Hessian矩阵为正定：$\mathbf{H}=\sum_{i=1}^{N}(diag(\boldsymbol{\mu}_i)-\boldsymbol{\mu}_i\boldsymbol{\mu}_i^T)\otimes\mathbf{x}_i\mathbf{x}_i^T$

► 小结：Softmax分类器Logistic回归

- Softmax分类器能实现多类分类，是对Logistic回归在两类分类任务上的扩展

- 优化算法和正则与两类分类Logistic回归类似。

CSDN
不 止 于 代 码

THANK YOU

AI100