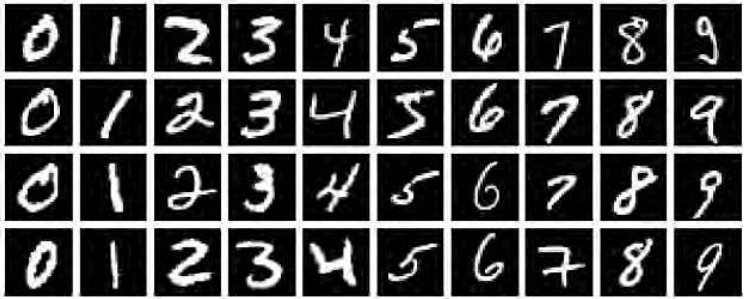

tensorflow / tensorflow
TensorFlow: Open sou
Clone this repo:
git clone https://tensor
Branches
master

11.5 多分类问题

逻辑回归：二分类问题

多分类问题：把输入样本划分为多个类别



自然顺序码

0——山鸢尾（Setosa）

1——变色鸢尾（Versicolour）

2——维吉尼亚鸢尾（Virginica）

独冷编码（one hot）

(0,1,1)——山鸢尾（Setosa）

(1,0,1)——变色鸢尾（Versicolour）

(1,1,0)——维吉尼亚鸢尾（Virginica）

独热编码（One-Hot Encoding）

- 使**非偏序关系**的数据，**取值不具有偏序性**
- 到原点等距

(1,0,0)——山鸢尾（Setosa）

(0,1,0)——变色鸢尾（Versicolour）

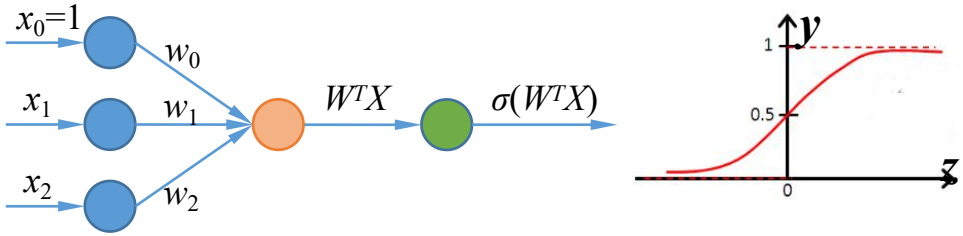
(0,0,1)——维吉尼亚鸢尾（Virginica）

应用

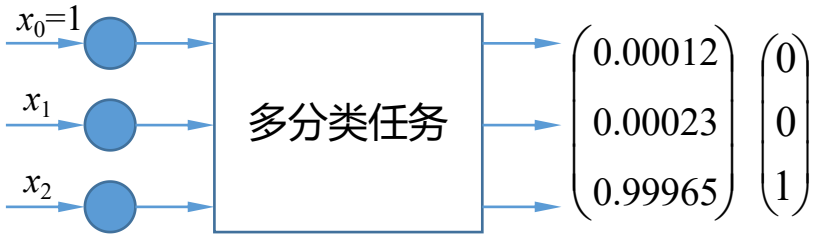
离散的**特征**

多分类问题中的**类别标签**

二分类



多分类

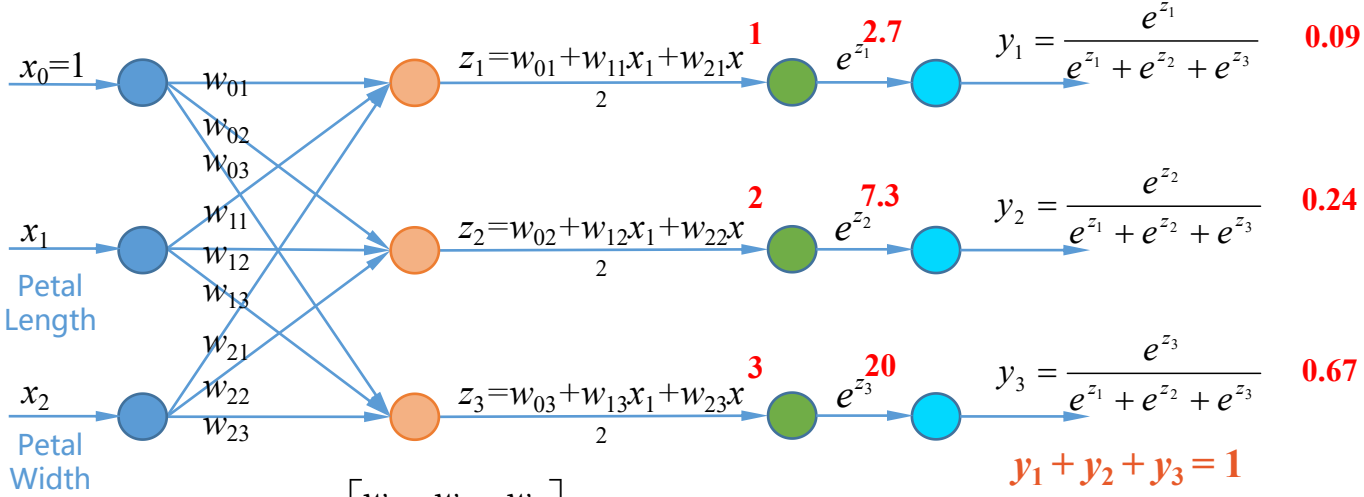


模型输出　标签

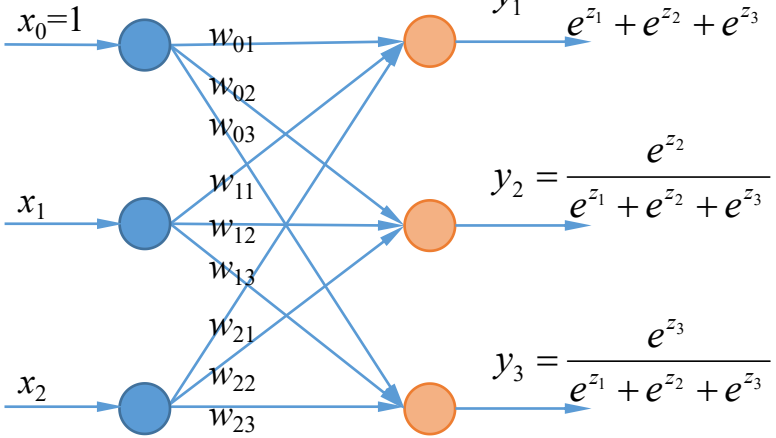
softmax()函数 $Y = soft\max(W^T X)$

广义线性回归, 实现多分类

例：使用属性**花瓣长度**和**花瓣宽度**，构造分类器，能够**识别3种类型**的鸢尾花

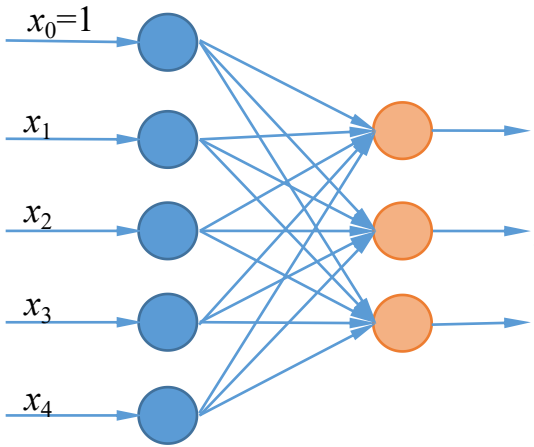


$$W=\begin{bmatrix} w_{01} & w_{02} & w_{03} \\ w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix}$$



$$W = \begin{bmatrix} w_{01} & w_{02} & w_{03} \\ w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix}$$

Softmax函数是Logistic函数在**多分类**问题上的推广



$$y_1=\frac{e^{z_1}}{e^{z_1}+e^{z_2}+e^{z_3}}$$

$$y_2=\frac{e^{z_2}}{e^{z_1}+e^{z_2}+e^{z_3}}$$

$$y_3=\frac{e^{z_3}}{e^{z_1}+e^{z_2}+e^{z_3}}$$

$$y_k=\frac{e^{z_k}}{\sum_{p=1}^{C}e^{z_p}}$$

k：输出的索引

z_k：第k个输出接收到的所有输入的线性组合

C：类别总数

$$W=\begin{bmatrix} w_{01} & w_{02} & w_{03} \\ w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{bmatrix}$$

$$z_1=w_{01}+w_{11}x_1+w_{21}x_2+w_{31}x_3+w_{41}x_4$$

$$z_1=w_{02}+w_{12}x_1+w_{22}x_2+w_{32}x_3+w_{42}x_4$$

$$z_1=w_{03}+w_{13}x_1+w_{23}x_2+w_{33}x_3+w_{43}x_4$$

- **二元交叉熵损失函数** (BCE)

$$Loss = -\sum_{i=1}^{n}[y_i \ln \hat{y}_i + (1-y_i)\ln(1-\hat{y}_i)]$$

- **多分类交叉熵损失函数** (CCE)

$$Loss = -\sum_{i=1}^{n}\sum_{p=1}^{C} y_{i,p} \ln(\hat{y}_{i,p})$$

i：样本的索引

n：样本的总数

$y_{i,p}$：第i个样本属于第p类的标记值，0/1

$\hat{y}_{i,p}$：第i个样本属于第p类的预测概率

C：类别总数

$$y_k = \frac{e^{z_k}}{\sum_{p=1}^{C} e^{z_p}}$$

$$Loss = -\sum_{i=1}^{n}\sum_{p=1}^{C} y_{i,p} \ln(\hat{y}_{i,p})$$

	样本	标记	预测值	结果判断
模型A	样本1	0 0 1	0.3 0.3 0.4	正确
	样本2	0 1 0	0.3 0.4 0.3	正确
	样本3	1 0 0	0.1 0.2 0.7	错误
模型B	样本1	0 0 1	0.1 0.2 0.7	正确
	样本2	0 1 0	0.1 0.7 0.2	正确
	样本3	1 0 0	0.3 0.4 0.3	错误

准确率：

模型A=模型B= 2/3 = 66.7%

交叉熵损失：

模型A

样本1：-(0×ln0.3 + 0×ln0.3+1×ln0.4)= -ln0.4=0.9162...

样本2：-(0×ln0.3 + 1×ln0.4+0×ln0.3)= -ln0.4=0.9612...

样本3：-(1×ln0.1 + 0×ln0.2+0×ln0.7)= -ln0.1=2.3025...

交叉熵损失和：**5.9677...**

模型B

样本1：-(0×ln0.1 + 0×ln0.2+1×ln0.7)= -ln0.7=0.3566...

样本2：-(0×ln0.1 + 1×ln0.7+0×ln0.2)= -ln0.7=0.3566...

样本3：-(1×ln0.3 + 0×ln0.4+0×ln0.3)= -ln0.3=1.2039...

交叉熵损失和：**1.9173...**

- **互斥的多分类问题**：每个样本只能够属于**一个类别**

  鸢尾花的识别

  手写数字识别

- **非互斥的多分类问题**：一个样本可以同时属于**多个类别**

  构建**多个**一对多的**逻辑回归**

  标签：

  彩色图片

  包括人物的图片

  包括汽车的图片

  户外图片

  室内图片