

逻辑回归的原理和模型已经比较熟了，因此可能笔记会少一些

1.1 逻辑回归的介绍

逻辑回归（Logistic regression，简称LR）虽然其中带有“回归”两个字，但逻辑回归其实是一个**分类**模型，并且广泛应用于各个领域之中。虽然现在深度学习相对于这些传统方法更为火热，但实则这些传统方法由于其独特的优势依然广泛应用于各个领域。

而对于逻辑回归而且，最为突出的两点就是其**模型简单**和**模型的可解释性强**。

逻辑回归模型的优劣势：

- 优点：实现简单，易于理解和实现；计算代价不高，速度很快，存储资源低；
- 缺点：容易欠拟合，分类精度可能不高

```
9
10 ## 用逻辑回归模型拟合构造的数据集
11 lr_clf = lr_clf.fit(x_features, y_label) #其拟合方程为  $y=w_0+w_1*x_1+w_2*x_2$ 
```

Step3:模型参数查看

```
1 ## 查看其对应模型的w
2 print('the weight of Logistic Regression:',lr_clf.coef_)
3
4 ## 查看其对应模型的w0
5 print('the intercept(w0) of Logistic Regression:',lr_clf.intercept_)
```

```
the weight of Logistic Regression: [[0.73455784 0.69539712]]
```

```
the intercept(w0) of Logistic Regression: [-0.13139986]
```

Step4:数据和模型可视化

查看w0和w的方法还真的忘了，记录一下

```
1 ## 利用.info()查看数据的整体信息
2 iris_features.info()
```

```
<class 'pandas.core.frame.DataFrame'>

RangeIndex: 150 entries, 0 to 149

Data columns (total 4 columns):
 #   Column                Non-Null Count  Dtype  
---  -
 0   sepal length (cm)     150 non-null   float64
 1   sepal width (cm)      150 non-null   float64
 2   petal length (cm)     150 non-null   float64
 3   petal width (cm)      150 non-null   float64
dtypes: float64(4)

memory usage: 4.8 KB
```

一般都是用.head()来查看，做kaggle也是直接看介绍，记录一下.info()

逻辑回归 原理简介：

Logistic回归虽然名字里带“回归”，但是它实际上是一种分类方法，主要用于两分类问题（即输出只有两种，分别代表两个类别），所以利用了Logistic函数（或称为Sigmoid函数），函数形式为

$$\text{logit}(z) = \frac{1}{1 + e^{-z}}$$

对于模型的训练而言：实质上来说就是利用数据求解出对应的模型的特定的w。

从而得到一个针对于当前数据的特征逻辑回归模型。

而对于多分类而言，将多个二分类的逻辑回归组合，即可实现多分类。

其实个人感觉还是不要用逻辑回归进行多分类预测，使用sigmoid函数和概率进行预测决定了它更是适合二分类问题