<https://blog.csdn.net/weixin_38278334/article/details/82320307>

**sklearn——CountVectorizer详解**

<https://blog.csdn.net/aliceyangxi1987/article/details/80617649>

# SVM 的核函数选择和调参

<https://blog.csdn.net/hfutdog/article/details/88085878>

# sklearn计算准确率、精确率、召回率、F1 score

召回率是比率tp / (tp + fn)，其中tp是真正性的数量，fn是假负性的数量. 召回率直观地说是分类器找到所有正样本的能力.

TP = sum(prediction==True AND groundtruth==True)

FP = sum(prediction==True AND groundtruth==False)

TN = sum(prediction==False AND groundtruth==False)

FN = sum(prediction==False AND groundtruth==True)