

# 凌风探梅的专栏

图像处理，图像分割，特征提取，机器学习，模式识别，深度学习等

目录视图 摘要视图 [RSS](#) 订阅

## 个人资料



凌风探梅



访问： 268119次

积分： 5012

等级： [BLOG > 6](#)

排名： 第4135名

原创： 61篇

转载： 1157篇

译文： 4篇

评论： 39条

## 联系方式

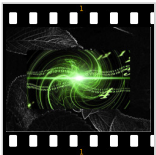
QQ群：  
1) OpenCV俱乐部  
186168905  
2) 视频/音频/图像/算法/ML  
群1：148111910

群2：157103105

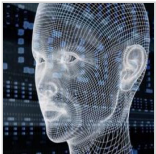
备注：加群需要回答问题，避免广告党。  
如果你是博客看到后加的，请注明“博客”并回答问题，只注明“博客”不回答问题的恕不加入。答案为和群相关的任何技术名词，不能出现1) 和2) 中的任何字眼

## 文章搜索

## 博客专栏



数字图像处理专题  
文章： 11篇  
阅读： 755



DeepID专栏  
文章： 6篇  
阅读： 309



OpenCV专栏  
文章： 16篇  
阅读： 2215

## 文章分类

## 准确率(Accuracy), 精确率(Precision), 召回率(Recall)和F1-Measure

2016-08-31 13:47 329人阅读 评论(0) 收藏 举报

分类： 科研指标 (3)

目录(?) [+]

## 准确率(Accuracy), 精确率(Precision), 召回率(Recall)和F1-Measure

yu Code 15 Comments

**机器学习**(ML),自然语言处理(NLP),信息检索(IR)等领域,评估(Evaluation)是一个必要的 工作,而其评价指标往往有如下几点:准确率(Accuracy),精确率(Precision),召回率(Recall)和F1-Measure。(注： 相对来说， IR 的 ground truth 很多时候是一个 Ordered List, 而不是一个 Bool 类型的 Unordered Collection，在都找到的情况下，排在第三名还是第四名损失并不是很大，而排在第一名和第一百名，虽然都是“找到了”，但是意义是不一样的，因此 更多可能适用于 **MAP** 之类评估指标。)

本文将简单介绍其中几个概念。中文中这几个评价指标翻译各有不同，所以一般情况下推荐使用英文。

现在我先假定一个具体场景作为例子。

假如某个班级有男生**80**人,女生**20**人,共计**100**人.目标是找出所有女生.

现在某人挑选出**50**个人,其中**20**人是女生,另外还错误的把30个男生也当作女生挑选出来了.

作为评估者的你需要来评估(**evaluation**)下他的工作

首先我们可以计算**准确率(accuracy)**,其定义是: 对于给定的测试数据集，分类器正确分类的样本数与总样本数之比。也就是损失函数是0-1损失时测试数据集上的准确率[1].

这样说听起来有点抽象，简单说就是，前面的场景中，实际情况是那个班级有男的和女的两类，某人(也就是定义中所说的分类器)他又把班级中的人分为男女两类。**accuracy**需要得到的是此**君分正确的人占总人数**的比例。很容易，我们可以得到:他把其中70(20女+50男)人判定正确了,而总人数是100人，所以它的**accuracy**就是70 %(70 / 100).

由准确率，我们的确可以在一些场合，从某种意义上得到一个分类器是否有效，但它并不总是能有效的评价一个分类器的工作。举个例子,google抓取了argcv 100个页面，而它索引中共有10,000,000个页面,随机抽一个页面，分类下,这是不是argcv的页面呢?如果以**accuracy**来判断我的工作，那我会把所有的页面都判断为"不是argcv的页面",因为我这样效率非常高(return false,一句话),而**accuracy**已经到了99.999%(9,999,900/10,000,000),完爆其它很多分类器辛辛苦苦算的值,而我这个算法显然不是需求期待的,那怎么解决呢?这就是**precision,recall**和**f1-measure**出场的时间了.

在说**precision,recall**和**f1-measure**之前,我们需要先需要定义TP,FN,FP,TN四种分类情况.

按照前面例子,我们需要从一个班级中的人中寻找所有**女生**,如果把这个任务当成一个分类器的话,那么女生就是我们需要的,而男生不是,所以我们称女生为"正类",而男生为"负类".

	相关(Relevant),正类	无关(NonRelevant),负类
被检索到 (Retrieved)	true positives(TP 正类判定为正类,例子中就是正确的判定"这位是女生")	false positives(FP 负类判定为正类,"存伪",例子中就是分明是男生却判断为女生,当下伪娘横行,这个错常有人犯)
未被检索到 (Not	false negatives(FN 正类判定为负类,"去真",例子中就是,分明是女生,这哥们却判断为男生--梁山伯同	true negatives(TN 负类判定为负类,也就是一个男生被判断为男生,像我这样的纯爷们一准儿就

- 并行计算专题 (9)
- 数字图像处理专题 (8)
- 视频处理专题 (2)
- OpenCV专题 (48)
- DeepID专题 (5)
- ImageSearch (26)
- ImageProcess (100)
- AuidoProcess (4)
- AudioFeature (1)
- DeepLearning (157)
- 机器学习 (78)
- 增强现实 (AR) (4)
- FFMpeg (15)
- ObjectDetect (30)
- 人脸检测 (34)
- 人脸识别 (31)
- 皮肤检测 (4)
- 性别识别 (0)
- C++ (32)
- C (29)
- CMake (0)
- Shell (17)
- OCR文字识别 (4)
- 实用工具 (20)
- CUDA (8)
- system (12)
- JSON (0)
- GearMan (3)
- Python (5)
- WSGI (2)
- CGI (1)
- 3D投影 (1)
- GPU显卡 (3)
- glibc (4)
- 手机卡 (1)
- 分布式 (1)
- zookeeper (2)
- 互联网 (4)
- 条形码 (1)
- 二维码 (1)
- 图像特征 (26)
- 图像去雾 (7)
- 图像边缘检测 (3)
- 距离度量 (2)
- 图像分析 (2)
- 概率论知识 (8)
- URL (22)
- 图像模糊 (6)
- 数学知识 (14)
- TF-IDF (3)
- 编辑器 (1)
- face morphing (1)
- Caffe (40)
- 硬盘维修 (1)
- 汇编 (1)
- 代码管理 (1)
- ZMQ (4)
- 立体匹配 (9)
- SpeechRecognition (3)
- AudioSearch (2)
- TextSearch (5)
- SearchEngineSoftware (2)
- 自动摘要 (2)
- Redis 数据库 (10)
- Clang (1)
- NoSQL (2)
- MongoDB (5)
- H265 (2)
- 管理 (2)
- 编解码知识 (30)

Retrieved)	学犯的错就是这个)	会在此处)
------------	-----------	-------

通过这张表,我们可以很容易得到这几个值:

TP=20

FP=30

FN=0

TN=50

**精确率(precision)**的公式是 $P = \frac{TP}{TP+FP}$ ,它计算的是所有"正确被检索的item(TP)"占有所有"实际被检索到的(TP+FP)"的比例.

在例子中就是希望知道此君得到的所有人中,正确的人(也就是女生)占有的比例.所以其precision也就是40%(20女生/(20女生+30误判为女生的男生)).

**召回率(recall)**的公式是 $R = \frac{TP}{TP+FN}$ ,它计算的是所有"正确被检索的item(TP)"占有所有"应该检索到的item(TP+FN)"的比例。

在例子中就是希望知道此君得到的女生占本班中所有女生的比例,所以其recall也就是100%(20女生/(20女生+ 0 误判为男生的女生))

F1值就是精确值和召回率的调和均值,也就是

$$\frac{2}{F_1} = \frac{1}{P} + \frac{1}{R}$$

调整下也就是

$$F_1 = \frac{2PR}{P+R} = \frac{2TP}{2TP+FP+FN}$$

例子中 F1-measure 也就是约为 57.143%( $\frac{2*0.4*1}{0.4+1}$ ).

需要说明的是,有人[2]列了这样个公式

$$F_a = \frac{(a^2+1)PR}{a^2(P+R)}$$

将F-measure一般化.

F1-measure认为精确率和召回率的权重是一样的,但有些场景下,我们可能认为精确率会更加重要,调整参数a,使用Fa-measure可以帮助我们更好的evaluate结果.

话虽然很多,其实实现非常轻松,点击[此处](#)可以看到我的一个简单的实现.

## References

[1] 李航. 统计学习方法[M]. 北京:清华大学出版社,2012.

[2] 准确率（Precision）、召回率（Recall）以及综合评价指标（F1-Measure）

=====

自己理解 + 我老师的说法就是，准确率就是**找得对**，召回率就是**找得全**。

大概就是你问问一个模型，这堆东西是不是某个类的时候，准确率就是 **它说是，这东西就确实是**的概率吧，召回率就是， **它说是，但它漏说了（1-召回率）**这么多。

=====

在信息检索、分类体系中，有一系列的指标，搞清楚这些指标对于评价检索和分类性能非常重要，因此最近根据网友的博客做了一个汇总。

### 准确率、召回率、F1

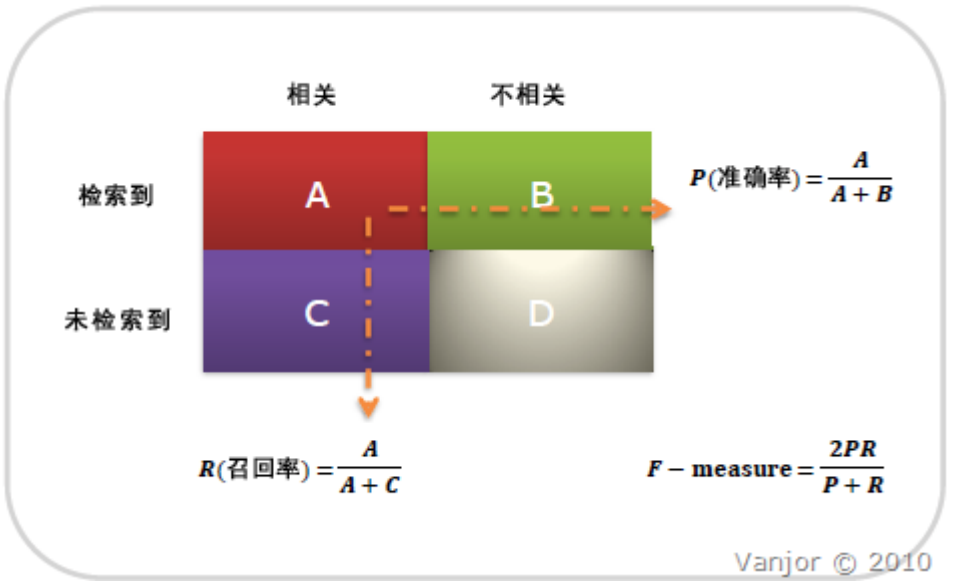
信息检索、分类、识别、翻译等领域两个最基本指标是**召回率(Recall Rate)**和**准确率(Precision Rate)**，召回率也叫查全率，准确率也叫查准率，概念公式：

召回率(**Recall**) = 系统检索到的相关文件 / 系统所有相关的文件总数

准确率(**Precision**) = 系统检索到的相关文件 / 系统所有检索到的文件总数

图示表示如下：

- 日志库 (1)
- Centos (2)
- 数据库 (8)
- Makefile (2)
- 流媒体 (35)
- 自然语言处理 (3)
- 数据挖掘 (3)
- 推荐系统 (23)
- valgrind (1)
- inotify (1)
- SQLite (1)
- GPUImage (11)
- WebP (3)
- 图像滤镜 (37)
- 颜色空间 (7)
- 数据结构 (1)
- 图像拼接 (19)
- Charles (1)
- 排序算法 (7)
- GLSL (13)
- Android (13)
- Spark (4)
- s-plus (1)
- KNN (4)
- RANSAC (1)
- 垃圾文本过滤 (4)
- IOS (4)
- 图像分割 (3)
- ImageNet (3)
- 敏感词 (8)
- Socket (6)
- 科研指标 (4)
- word2vec (1)
- 中文分词 (6)
- Weka (1)
- 贝叶斯分类器 (1)
- LSTM (2)
- 图像质量评价 (6)
- 模糊检测 (1)
- nginx (10)
- PHP (0)
- 图像增强 (3)
- SIMD (6)
- 手背静脉识别 (1)
- 系统设计 (5)
- alphago (2)
- 视频分割 (1)
- scala (1)
- Docker (8)
- HDR (3)
- Android (4)
- 虚拟现实 (7)
- 大数据 (3)
- TensorFlow (7)
- 测距 (5)
- 立体视觉 (1)
- FFT (3)
- QT (1)
- SSE (1)
- Elasticsearch (3)
- 服务器架构 (4)
- LBS (13)
- 开源项目 (2)
- 前背景分离 (2)
- ADAS相关 (67)
- AR (3)
- Windows (2)
- Object-C (1)
- 魔方 (2)
- 显著图 (2)
- 手势识别 (2)





数码相机 (1)  
词袋模型BOW (3)  
LDA (1)  
视频特征 (1)  
视频音频处理 (3)  
Lua (1)  
双目立体视觉 (1)  
wurenjiashi (1)  
graphics rendering engines (1)  
Git (20)  
自动驾驶 (3)  
聊天机器人 (1)  
色情识别 (1)  
FPGA (1)  
机器人 (1)  
RabbitMQ (1)  
SVM (1)  
ARM (7)  
工作生活 (12)  
fisheye (7)  
OpenMP (1)  
Face2Face-Reenactment (5)  
Axure (1)  
容器 (1)  
算法 (2)  
编程 (1)  
Paper (1)  
论文 (1)  
安全 (2)  
OpenStack (1)  
Linux (20)  
SUSE (2)  
汽车 (3)  
redmine (5)  
语音识别 (3)  
云直播 (1)  
Bundler (1)  
SLAM (4)  
Web系统 (2)  
Samba (6)  
串口调试 (2)  
无人驾驶 (2)  
相机校正 (6)  
传感器 (1)  
计算机病毒 (2)  
数据加密 (10)  
Unity3D (1)  
VR (4)  
MR (3)  
行人检测 (5)  
图像融合 (9)  
Eigen (1)  
微信开发 (1)  
图像压缩 (1)  
相机 (1)  
HTML (1)  
跟踪算法 (1)  
Vi (1)  
Vim (1)  
Svn (1)  
Dsp (1)  
线性代数 (1)  
嵌入式 (2)  
播放器 (10)  
OpenGL ES (0)

## 文章存档

2016年11月 (5)  
2016年10月 (70)  
2016年09月 (118)

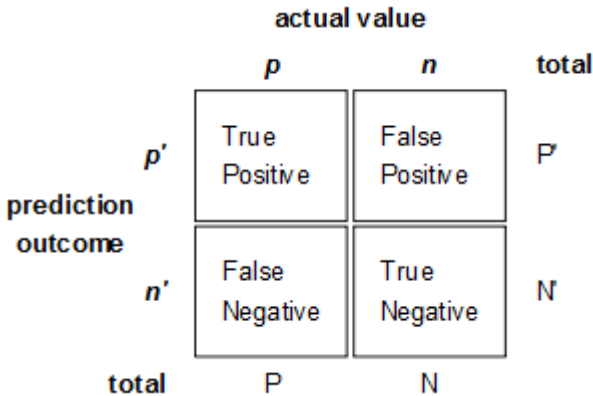
更加具体的，曲线与坐标轴之间的面积应当越大。

最理想的系统，其包含的面积应当是1，而所有系统的包含的面积都应当大于0。这就是用以评价信息检索系统的最常用性能指标，平均准确率mAP其规范的定义如下:(其中P，R分别为准确率与召回率)

$$mAP = \int_0^1 P(R) dR$$

## ROC和AUC

ROC和AUC是评价分类器的指标，上面第一个图的ABCD仍然使用，只是需要稍微变换。



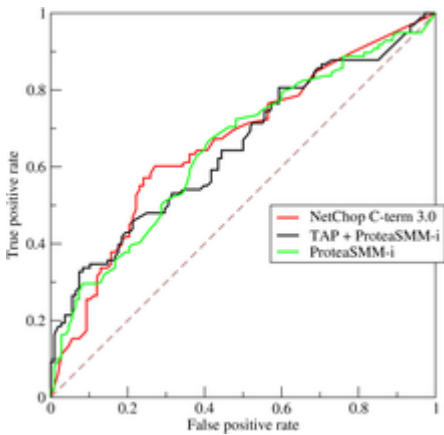
回到ROC上来，ROC的全名叫做Receiver Operating Characteristic。

## ROC关注两个指标

True Positive Rate ( TPR ) =  $TP / [ TP + FN ]$  , TPR代表能将正例分对的概率

False Positive Rate( FPR ) =  $FP / [FP + TN]$  , FPR代表将负例错分为正例的概率

在ROC空间中，每个点的横坐标是FPR，纵坐标是TPR，这也就描绘了分类器在TP（真正的正例）和FP（错误的正例）间的trade-off。ROC的主要分析工具是一个画在ROC空间的曲线——ROC curve。我们知道，对于二值分类问题，实例的值往往是连续值，我们通过设定一个阈值，将实例分类到正类或者负类（比如大于阈值划分为正类）。因此我们可以变化阈值，根据不同的阈值进行分类，根据分类结果计算得到ROC空间中相应的点，连接这些点就形成ROC curve。ROC curve经过 (0,0) (1,1)，实际上(0, 0)和(1, 1)连线形成的ROC curve实际上代表的是一个随机分类器。一般情况下，这个曲线都应该处于(0, 0)和(1, 1)连线的上方。如图所示。



用ROC curve来表示分类器的performance很直观好用。可是，人们总是希望能有一个数值来标志分类器的好坏。于是**Area Under roc Curve(AUC)**就出现了。顾名思义，AUC的值就是处于ROC curve下方的那部分面积的大小。通常，AUC的值介于0.5到1.0之间，较大的AUC代表了较好的Performance。

### AUC计算工具：

<http://mark.goadrich.com/programs/AUC/>

P/R和ROC是两个不同的评价指标和计算方式，一般情况下，检索用前者，分类、识别等用后者。

参考链接：

<http://www.vanjoj.org/blog/2010/11/recall-precision/>

<http://bubblexc.com/y2011/148/>

<http://wenku.baidu.com/view/ef91f011cc7931b765ce15ec.html>

: Recall, 又称“查全率”——还是查全率好记, 也更能体现其实质意义。

## 准确率

“召回率”与“准确率”虽然没有必然的关系（从上面公式中可以看到），在实际应用中，是相互制约的。要根据实际需求，找到一个平衡点。

当我们问检索系统某一件事的所有细节时（输入检索query查询词），Recall指：检索系统能“回忆”起那些事的多少细节，通俗来讲就是“回忆的能力”。“能回忆起来的细节数”除以“系统知道这件事的所有细节”，就是“记忆率”，也就是recall——召回率。简单的，也可以理解为查全率。

顶  
1

0

阅读排行

NVIDIA GPU 运算能力列

(4009)

Caffe 编译安装

(3258)

ZMQ guider

(1847)

DeepLearning tutorial (转)

(1440)

YOLO: Real-Time Objec

(1402)

深度学习算法简介

(1368)

Augmented reality

(1286)

人工智能中对机器学习非常

(1200)

用SRS和FFMPEG进行直播

(1161)

CNN笔记：通俗理解卷积

(1142)

评论排行

立体匹配算法实现之：AdaptWei

(12)

Android Studio使用OpenGL ES 3.0

(3)

摄像头图像桶形畸变校正测试

(2)

DeepLearning tutorial (转)

(2)

几个视频中行为识别的底层实现

(2)

有趣的机器学习：最简明易懂

(2)

opencv实现图像的拼接功能

(2)

在centos6.5中安装 GitLab

(2)

视频压缩编码和音频压缩编码

(1)

【Caffe实践】基于Caffe的行人检测

(1)

推荐文章

\* 2016 年最受欢迎的编程语言是什么？

\* Chromium扩展（Extension）的页面（Page）加载过程分析

\* Android Studio 2.2 来啦

\* 手把手教你做音乐播放器（二）技术原理与框架设计

\* JVM 性能调优实战之：使用阿里开源工具 TProfiler 在海量业务代码中精确定位性能代码

最新评论

opencv实现图像的拼接功能

凌风探梅: 你是 3 2 位的系统么？拼起来超过 4 G了 ？

opencv实现图像的拼接功能

TaidyPanda: 对于高分辨率的大图,内存会吃不消,有没有什么解决办法

GML C++ Camera Calibration Tool

Shy\_Big\_Boy: 博主，您好@请问用这个工具箱对相机进行标定，怎么获取相机的外部参数R,T？

立体匹配算法实现之：AdaptWei

guanluanjicai68209: 亲，我下载您后面给的代码是Region Growing滴，有没有ASW的代码呢，论文需要做对比实验， ...

摄像头图像桶形畸变校正测试

凌风探梅: 您是做什么的？可以留个Q Q

摄像头图像桶形畸变校正测试

ideallyworld200: 您好，我看了您转的这边文档，最近刚好也在学习这方面的，请问有复现文档结果吗？我复现的结果没文档中的原理...

聚焦和增强卷积神经网络

凌风探梅: 最老程序员闫涛文章的内容很好，但是文章的标题有点问题，文中讲述的是递归神经网络（RNN），而标题却是...

在centos6.5中安装 GitLab 全过程

凌风探梅: bitmami的一键安装工具可以完美解决你的问题，之前

我的同类文章

科研指标 (3)

• 准确率,召回率,F值,ROC,AUC

2016-08-31

阅读 74

• ROC曲线

2015-11-25

阅读 306

• ROC曲线 1

2015-11-25

阅读 153

参考知识库

阿里云机器学习算法应用实践

准确率Precision召回率Recall以及综合评价指标F1-Measure

统计机器学习入门——线性模型选择与正则化2

关于准确率Precision召回率RecallF1-Measure的解析

《C语言/C++学习指南》加密解密篇（安全相关算法）

准确率Precision召回率Recall以及综合评价指标F1-Measure

统计机器学习入门——线性模型选择与正则化1

准确率Precision召回率Recall以及综合评价指标F1-Measure

统计机器学习入门——分类2

准确率Precision召回率Recall以及综合评价指标F1-Measure

猜你在找

查看评论

暂无评论

您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

\* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

核心技术类目

全部主题

Hadoop

AWS

移动游戏

Java

Android

iOS

Swift

智能硬件

Docker

OpenStack

VPN

Spark

ERP

IE10

Eclipse

CRM

JavaScript

数据库

Ubuntu

NFC

WAP

BI

HTML5

Spring

Apache

.NET

API

HTML

SDK

IIS

Fedora

XML

LBS

Unity

Splashtop

UML

components

Windows Mobile

Rails

QEMU

KDE

Cassandra

CloudStack

FTC

coremail

OPhone

CouchBase

云计算

iOS6

Rackspace

Web App

SpringSide

Maemo

Compuware

大数据

apttech

Perl

Tornado

Ruby

Hibernate

ThinkPHP

HBase

Pure

Solr

Angular

Cloud Foundry

Redis

Scala

Django

Bootstrap

先把你安装全部卸掉，用一键安装工具安装，做好相...

在centos6.5中安装 GitLab 全过和  
yangeguang: !] There was an error parsing `Gemfile`: compile e...

【Caffe实践】基于Caffe的人脸ㄥ  
wuxue1991sdu: 您好，我想问一下为什么accuracy和loss层都是用euclidean loss呢？

公司简介 | 招贤纳士 | 广告服务 | 银行汇款帐号 | 联系方式 | 版权声明 | 法律顾问 | 问题报告 | 合作伙伴 | 论坛反馈

网站客服      杂志客服      微博客服      webmaster@csdn.net      400-600-2320 | 北京创新乐知信息技术有限公司 版权所有 | 江苏知之为计算机有限公司 |

江苏乐知网络技术有限公司

京 ICP 证 09002463 号 | Copyright © 1999-2016, CSDN.NET, All Rights Reserved 