



移民容易的国家



联系我们



请扫描二维码联系客服
webmaster@csdn.net
400-660-0108
QQ客服 客服论坛

关于 招聘 广告服务 网站地图
©2018 CSDN版权所有 京ICP证09002463号
百度提供搜索支持

经营性网站备案信息
网络110报警服务
中国互联网举报中心
北京互联网违法和不良信息举报中心



机器学习常用算法总结

2018年8月21日 23:26:48

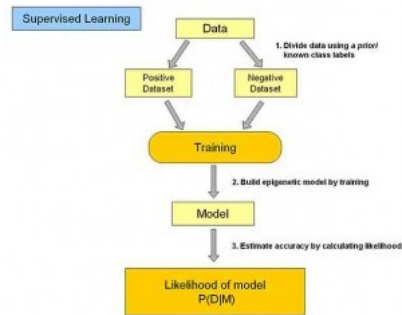


学习无疑是当前数据分析领域的一个热点内容。很多人在平时的工作中都或多或少会用到机器学习算法。本文旨在提供一个参考。机器学习的算法很多，很多算法是一类算法，而有些算法又是从其他算法中延伸出来的。本文主要介绍机器学习的方式，第二个方面是算法的类似性。

一、学习方式

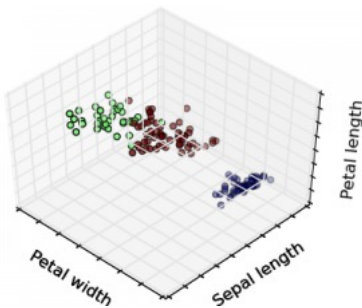
数据类型的不同，对一个问题的建模有不同的方式。在机器学习或者人工智能领域，人们首先会考虑数据是一个什么样的数据，这是一个不错的想法，这样可以让人们在建模和算法选择的时候考虑能根据输入数据来选择最合适的方式。机器学习主要有三种主要的学习方式：监督学习，非监督学习，半监督学习和强化学习。

1、监督式学习：



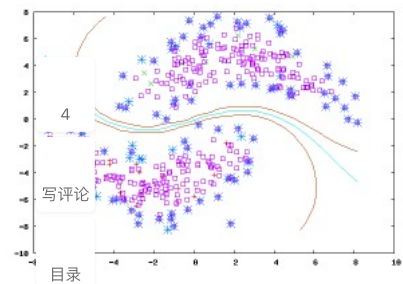
在监督式学习下，输入数据被称为“训练数据”，每组训练数据有一个明确的标识或结果，如对防垃圾邮件系统中“垃圾邮件”“非垃圾邮件”，对手写数字识别中的“1”，“2”，“3”，“4”等。在建立预测模型的时候，监督式学习建立一个学习过程，将预测结果与“实际结果”的实际结果进行比较，不断的调整预测模型，直到模型的预测结果达到一个预期的准确率。监督式学习的常见应用场景如分类问题和回归问题。常见监督式学习算法有决策树学习(ID3,C4.5等)，朴素贝叶斯分类，最小二乘回归，逻辑回归（Logistic Regression），支撑向量机，以及反向传递神经网络（Back Propagation Neural Network）等等。

2、非监督式学习：



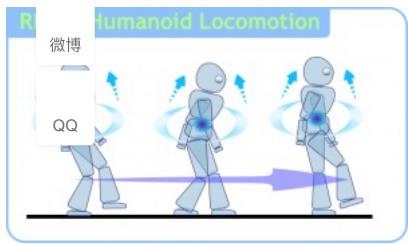
在非监督式学习中，数据并不被特别标识，学习模型是为了推断出数据的一些内在结构。常见的应用场景包括关联规则的学习以及聚类分析。常见非监督学习算法包括奇异值分解、主成分分析，独立成分分析，Apriori算法以及k-Means算法等等。

3、半监督式学习：



化学习方式下，输入数据部分被标识，部分没有被标识，这种学习模型可以用来进行预测，但是模型首先需要学习数据的内在结构。应用场景包括分类和回归，算法包括一些对常用监督式学习算法的延伸，这些算法首先试图对未标识数据进行预测，然后再对标识的数据进行预测。如图论推理算法（Graph Inference）或者拉普拉斯支持向量机（Laplacian SVM）等。

4、 学习：

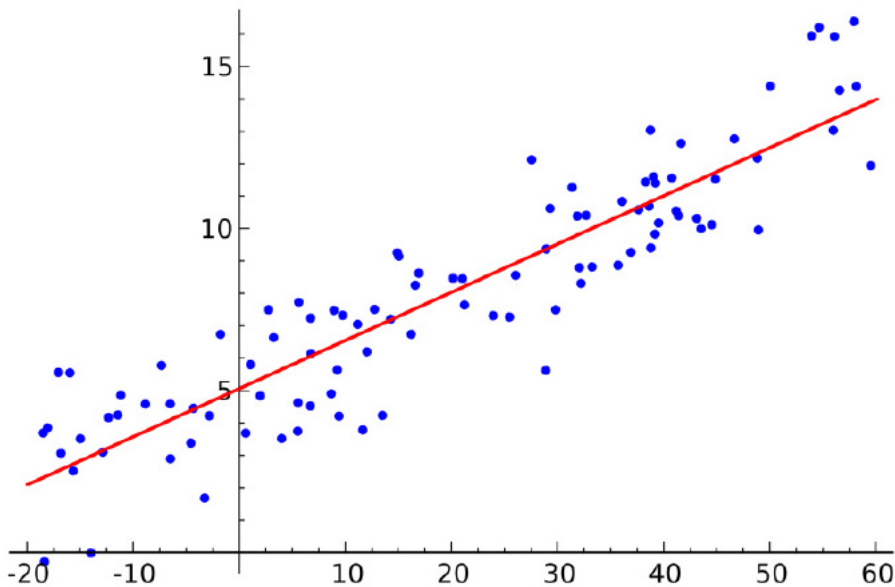


在这种学习模式下，输入数据作为对模型的反馈，不像监督模型那样，输入数据仅仅是作为一个检查模型对错的方式，在强化学习中数据直接反馈到模型，模型必须对此立刻作出调整。常见的应用场景包括动态系统以及机器人控制等。常见算法包括Q-Learning以及时间差分学习（Temporal difference learning）。在企业数据应用的场景下，人们最常用的可能就是监督式学习和非监督式学习的模型。在图像识别域，由于存在大量的非标识的数据和少量的可标识数据，目前半监督式学习是一个很热的话题。而强化学习更多的应用在机器人控制及非进行系统控制的领域。

二、 算法类似性

根据算法的功能和形式的类似性，我们可以把算法分类，比如说基于树的算法，基于神经网络的算法等等。当然，机器学习的范围非常庞大，有些算法很难明确归类到某一类。有些分类来说，同一分类的算法可以针对不同类型的问题。这里，我们尽量把常用的算法按照最容易理解的方式进行分类。

1、 回归算法：



回归算法是试图采用对误差的衡量来探索变量之间的关系的一类算法。回归算法是统计机器学习的利器。在机器学习领域，人们说起回归，有时候是指一类问题，有时候是指一类算法，这一点常常会使初学者有所困惑。常见的回归算法包括：最小二乘法（Ordinary Least Square），逻辑回归（Logistic Regression），逐步式回归（Stepwise Regression），多元自适应回归样条（Multivariate Adaptive Regression Splines）以及本地散点平滑估计（Locally Estimated Scatterplot Smoothing）。

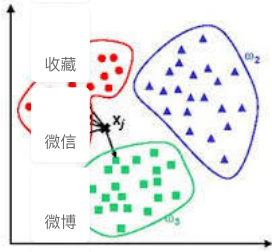
通常，回归可以被用于在现实世界的应用，如：

- 信用评分
- 营销活动成功率
- 某产品的收入
- 特定的日子里会发生地震吗？

写评论

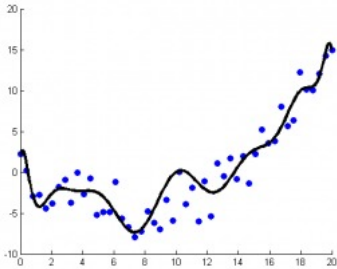
2、实例的算法

目录



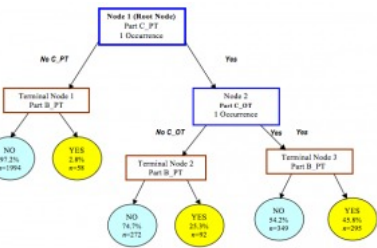
实例的算法常常用来对决策问题建立模型，这样的模型常常先选取一批样本数据，然后根据某些近似性把新数据与样本数据进行比较。通过这种匹配。因此，基于实例的算法常常也被称为“赢家通吃”学习或者“基于记忆的学习”。常见的算法包括 k-Nearest Neighbor(KNN), 学习矢量量 (Learning Vector Quantization, LVQ)，以及自组织映射算法 (Self-Organizing Map , SOM)。

3、正则化方法



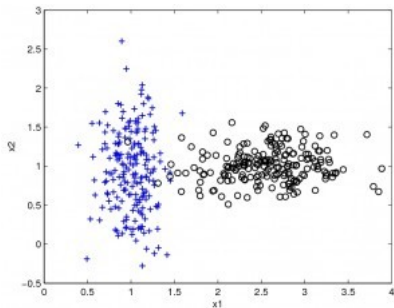
正则化方法是其他算法（通常是回归算法）的延伸，根据算法的复杂度对算法进行调整。正则化方法通常对简单模型予以奖励而对复杂算法予以惩罚。算法包括：Ridge Regression , Least Absolute Shrinkage and Selection Operator (LASSO)，以及弹性网络 (Elastic Net)。

4、决策树学习



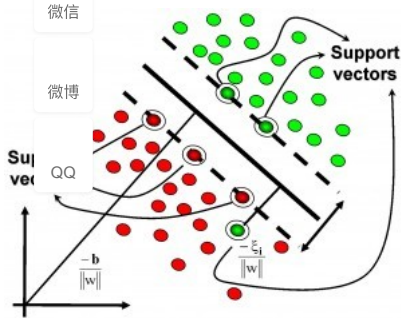
决策树算法根据数据的属性采用树状结构建立决策模型，决策树模型常常用来解决分类和回归问题。常见的算法包括：分类及回归树 (Classification regression Tree , CART)，ID3 (Iterative Dichotomiser 3), C4.5, Chi-squared Automatic Interaction Detection(CHAD), Decision Stump, 随机森林 (Random Forest)，多元自适应回归样条 (MARS) 以及梯度推进机 (Gradient Boosting Machine , GBM)。

5、贝叶斯方法



- 4 现实中的例子：
- 一个电子邮件为垃圾邮件或非垃圾邮件
 - 写评论的文章分为技术类、政治类或体育类
 - 一段文字表达积极的情绪，或消极的情绪？
 - 人脸识别软件

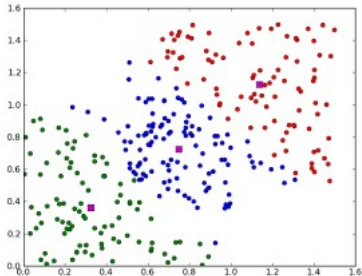
6、收藏 核的算法



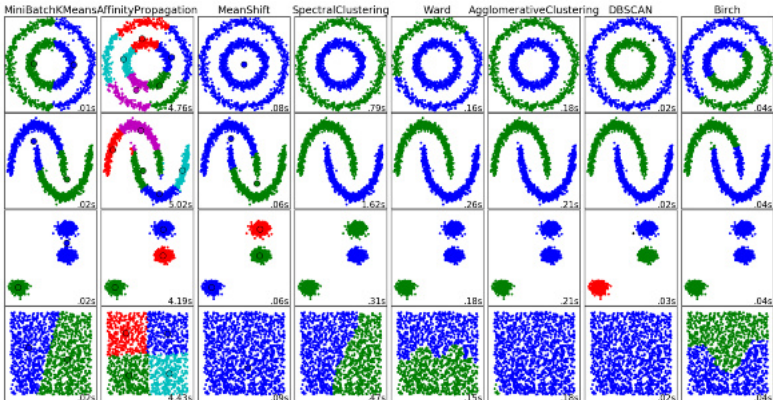
基于核的算法中最著名的莫过于支持向量机（SVM）了。基于核的算法把输入数据映射到一个高阶的向量空间，在这些高阶向量空间里，有些分类问题能够更容易的解决。常见的基于核的算法包括：支持向量机（Support Vector Machine，SVM），径向基函数（Radial Basis Function，RBF），判别分析（Linear Discriminate Analysis，LDA）等。

就规模而言，其中一些最主要的问题已经使用支持向量机解决了（通过适当的修改），如，入广告显示，人类的剪接位点识别，基于图像的性别检测图像分类等等。

7、聚类算法



聚类，就像回归一样，有时候人们描述的是一类问题，有时候描述的是一类算法。聚类算法通常按照中心点或者分层的方式对输入数据进行归并。所算法都试图找到数据的内在结构，以便按照最大的共同点将数据进行归类。常见的聚类算法包括 k-Means算法以及期望最大化算法（Expectation Maximization，EM）。



每一种聚类算法都不太一样，这里有一些：

- 质心的算法
- 4 连通性的算法
- 密度的算法
- 写评论
- 目录
- 网络/深度学习

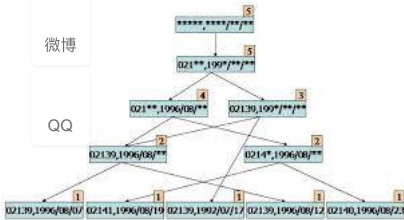
收藏

8、关联规则学习

微信

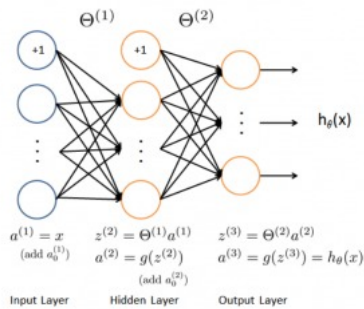
微博

QQ



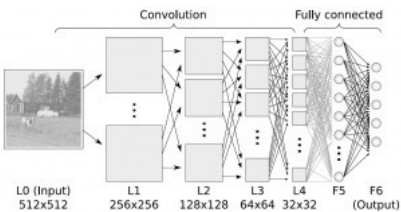
关联规则学习通过寻找最能够解释数据变量之间关系的规则，来找出大量多元数据集中有用的关联规则。常见算法包括 Apriori算法和Eclat算法等。

9、人工神经网络



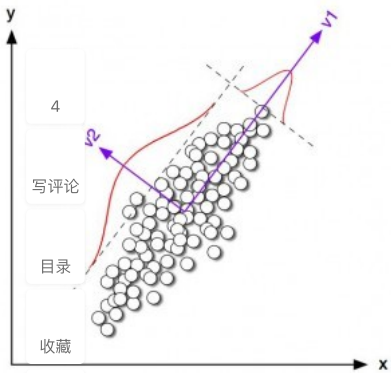
人工神经网络算法模拟生物神经网络，是一类模式匹配算法。通常用于解决分类和回归问题。人工神经网络是机器学习的一个庞大的分支，有几百种方法。（其中深度学习就是其中的一类算法，我们会单独讨论），重要的人工神经网络算法包括：感知器神经网络（Perceptron Neural Network），反向传播（Propagation），Hopfield网络，自组织映射（Self-Organizing Map, SOM）。学习矢量量化（Learning Vector Quantization, LVQ）。

10、深度学习

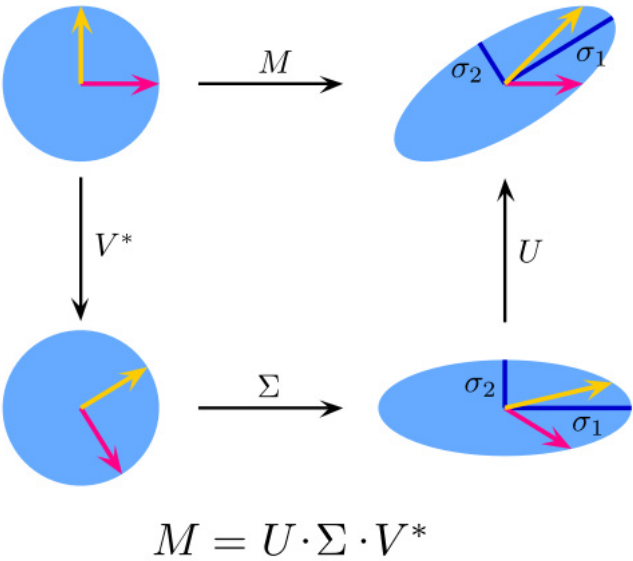


深度学习算法是对人工神经网络的发展。在近期赢得了很多关注，特别是[百度也开始发力深度学习后](#)，更是在国内引起了很多关注。在计算能力变价的今天，深度学习试图建立大得多也复杂得多的神经网络。很多深度学习的算法是半监督式学习算法，用来处理存在少量未标识数据的大数据集。常见的算法包括：受限波尔兹曼机（Restricted Boltzmann Machine, RBN），Deep Belief Networks（DBN），卷积网络（Convolutional Network），堆编码器（Stacked Auto-encoders）。

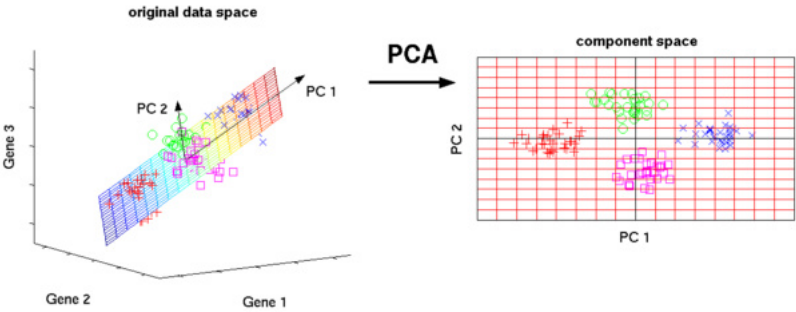
11、降低维度算法



类算法一样，降低维度算法试图分析数据的内在结构，不过降低维度算法是以非监督学习的方式试图利用较少的信息来归纳或者解释数据。这类算法通常用于数据的可视化或者用来简化数据以便监督式学习使用。常见的算法包括：主成份分析（ Principle Component Analysis， PCA ），偏最小二乘回归（ Partial Least Square Regression， PLS ）， Sammon映射，多维尺度（ Multi-Dimensional Scaling, MDS ），投影追踪（ Projection Pursuit ）等。

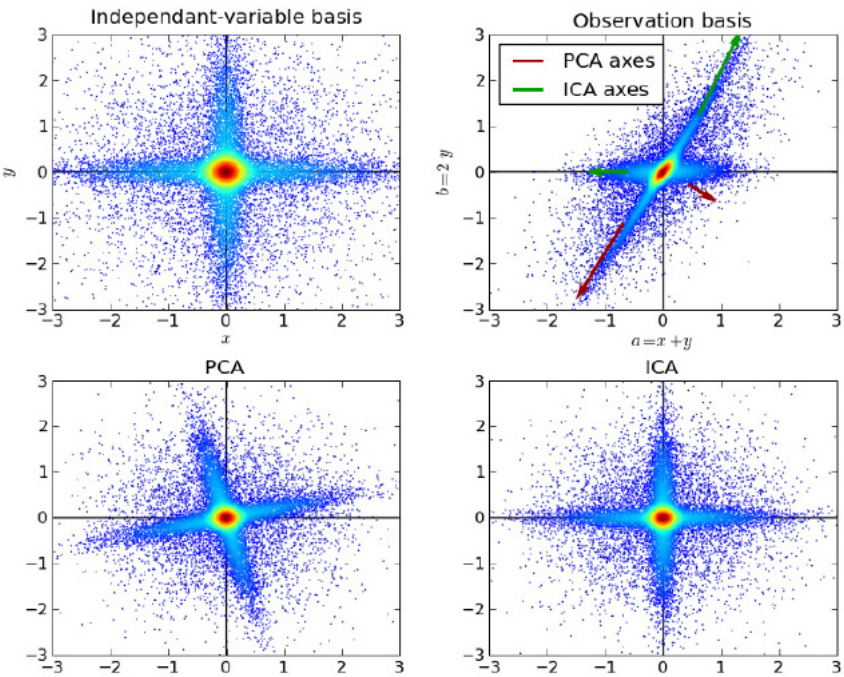


奇异值分解示意图



PCA降维示意图

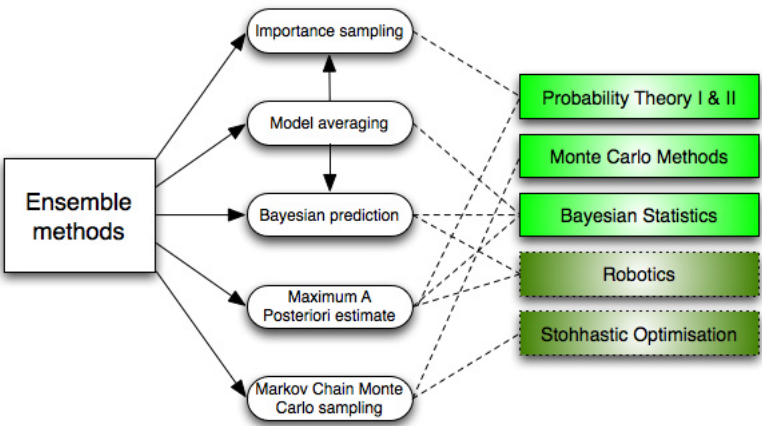
- 4
- 写评论
- 目录
- 收藏
- 微信
- 微博
- QQ



独立成分分析示意图

其中，ICA 和 PCA 是相关的，但是它是一种更强大的技术，当那些经典的方法完全失效的时候，它能够从数据源中发现潜在的因素。它的应用包括数字文档数据库，经济指标和心理测量。

12、集成算法：



集成算法用一些相对较弱的学习模型独立地就同样的样本进行训练，然后把结果整合起来进行整体预测。集成算法的主要难点在于究竟集成哪些独立学习模型以及如何把学习结果整合起来。这是一类非常强大的算法，同时也非常流行。常见的算法包括：Boosting， Bootstrapped Aggregation (Bagging)， AdaBoost， 堆叠泛化 (Stacked Generalization， Blending)， 梯度推进机 (Gradient Boosting Machine, GBM)， 随机森林 (Random Forest)。

那么集成方法是怎样工作的，为什么他们会优于单个的模型？

- 他们拉平了输出偏差：如果你将具有民主党倾向的民意调查和具有共和党倾向的民意调查取平均，你将得到一个中和的没有倾向一方的结果。
- 它们减小了方差：一堆模型的聚合结果和单一模型的结果相比具有更少的噪声。在金融领域，这被称为多元化——多只股票的混合投资要比一只股票变化更小。这就是为什么越多你的模型会越好，而不是数据点越少越好。
- 它们不太可能产生过拟合：如果你有一个单独的没有过拟合的模型，你是用一种简单的方式（平均，加权平均，逻辑回归）将这些预测结果结合起来，然后就没有产生过拟合了。

参考资料：

机器学习常见算法分类汇总

ML 工程师需了解的 10 大算法

机器学习10大经典算法

4 本文为博主原创文章，未经博主允许不得转载。https://blog.csdn.net/JIEJINQUANIL/article/details/52270178

文章 机器学习 算法 ▼ 查看关于本篇文章更多信息

写评论

上一篇 翻转子串的解题技巧

下一篇 一些在线编程练习或调试网站

目录

区块链以太坊开发人才稀少？薪资高到爆！只因太简单没人学

区块链收藏以太坊DApp开发是怎么炼成的？区块链的日益火爆和备受追捧，使得区块链开发人才稀缺，同时更加伴随着高薪，想要转型却不了解，特整理出学习大纲

微信

想对作者说点什么？ 我来说两句

微博

机器学习应该知道的3种机器学习算法

QQ 798

原文：3 machine Learning Algorithms You Need to Know 作者：Eleni Markou 翻译：Vincent 译者注：决策...

简单易学的机器学习算法——决策树之ID3算法

4.9万

一、决策树分类算法概述 二、ID3

机器学习常用算法总结 - CSDN博客

机器学习的算法很多,很多算法是一类算法,而有些算法...因此,基于实例的算法常常也被称为“赢家通吃”学习...

机器学习之各种算法(1) - CSDN博客

当然,机器学习的范围非常庞大,有些算法很难明确归类到某一类。而对于有些分类来...因此,基于实例的算法常常...

服了！人工智能应届生平均年薪30W只是“白菜价”

机器学习|深度学习|图像处理|自然语言处理|无人驾驶，这些技术都会吗？看看真正的人工智能师...

机器学习算法集锦

5209

机器学习机器学习(Machine Learning, ML)是一门多领域交叉学科，涉及概率论、统计学、逼近...

younger chi

机器学习是一种能从训练数据当中学习模型的一种软件,它还能把这些模型运用到新...由于赢家通吃的心理,像谷...

机器学习之各种算法 - CSDN博客

当然,机器学习的范围非常庞大,有些算法很难明确归类到某一类。而对于有些分类来...因此,基于实例的算法常常...

上手机器学习，从搞懂这十大经典算法开始

634

2 - 逻辑回归 逻辑回归是机器学习借自统计领域的另一项技术，用于解决二元分类问题（有两个...

几个常用机器学习算法 - 逻辑回归

910

1 回归在数学上来说,回归是给定一个点集，然后用一条曲线去拟合。如果这条曲线是一条直线...

【机器学习】机器学习算法之旅 - CSDN博客

在这篇文章中,我们将介绍最流行的机器学习算法。浏览该领域的主要算法以了解可用...基于这个原因,基于实例...

今天看到一篇不错的[机器学习入门贴 - Zhang_Yu_Joseph](#)的博客 - ...

机器学习入门 机器学习无疑是当前数据分析领域的一个热点内容。很多人在平时的...因此,基于实例的算法常常...

[机器学习十大算法,完整集合](#)

下载 2018年08月08日

写评论 *版本: 只允许上传png/jpeg/jpg/gif格式的图片,且小于3M *详细原因: 取消 提交 机器学习十大...

老...从吃了这个，每晚要搞我好几次！

泰斗 收藏 顶新

我：[机器学习 机器学习干货站- 资源 - CSDN博客](#)

我...机器学习 机器学习干货站跳至正文机器学习 Deep Learning 所有主题关注我们 ...因此,基于实例的算法常常...

大...结构详解:从数据获取到深度学习(内含福利) - CSDN博客

↑ 点 QQ 蓝字关注我们,和小伙伴一起聊技术!机器学习(Machine Learning,ML)是一门多领域交叉学科,涉及概率...

2017-我的秋招之路：[机器学习/算法工程师](#)(含大量面经)

6582

前言：从今年的8月1号正式开启秋招到现在，过去三个多月了，发生了很多事情，经历了N多场笔试和面试， ...

[机器学习--归纳总结](#)

1579

假设空间 归纳与演绎 布尔概念学习 版本空间 归纳偏好 线性回归 基本形式 线性回归 最小二乘参...

[机器学习:算法介绍 - Mojoooo-jie - CSDN博客](#)

这篇文章介绍几个最流行的机器学习算法。现在有很多机器学习算法,困难的是进行...出于这个原因,基于实例的...

一 [机器学习分类](#) - CSDN博客

机器学习是一种能够赋予机器学习的能力以此让它完成直接编程无法完成的功能的方法...因此,基于实例的算法常...

[机器学习面试要点总结](#)

155

1、说到xgboost，不得不说gbdt。gbdt详细说明见下面这篇文章。 <https://www.cnblogs.com/ModifyRong/p/7...>

[机器学习面试常用算法知识点梳理总结](#)

1576

原文地址：<http://www.cnblogs.com/tornadomeet/p/3395593.html> 前言： 找工作时 (l...

[机器学习常见算法个人总结（面试用）](#)

1362

朴素贝叶斯 参考[1] 事件A和B同时发生的概率为在A发生的情况下发生B或者在B发生的情况下发...

[机器学习的9个基础概念和10种基本算法总结](#)

3158

1.基础概念： (1) 10折交叉验证：英文名是10-fold cross-validation，用来测试算法的准确性...

[图解十大经典机器学习算法入门](#)

4617

弱人工智能近几年取得了重大突破，悄然间，已经成为每个人生活中必不可少的一部分。以我们...

[\[下载CAMBLY\]随时随地学英语](#)

随时随地学习，把欧美外教装在口袋里！现在注册可领取15分钟免费课程！

8/9/2018

机器学习常用算法总结 - CSDN博客

干货 | 上手机器学习，从搞懂这十大经典算法开始

翻译

技大本营 (rgznai100)

参与 | 林椿昀 编辑 | 波波, Donna

在机器学习领域, “没有免...

4

机器学习 (常见面试之机器学习算法思想简单梳理)

sup

写评论 |

(classification[k nearest neighbor、decision tree、naive bayes、logistic regression...

5128

成为机器学习算法工程师，你需要这些必备技能

成为

格的开发工程师不是一件简单的事情, 需要掌握从开发到调试到优化等一系列能力, ...

收藏

3137

8种机器学习算法比较

简介

学习算法太多了, 分类、回归、聚类、推荐、图像识别领域等等, 要想找到一个合适算...

微信

微博

1.9万

5大机器学习算法总结

算法

QQ

动态规划算法 一、基本概念 动态规划过程是: 每次决策依赖于当前状态, 又随即引起状态的转移...

2471

oa办公管理系统

百度广告



机器学习十大算法

2017年12月05日 4.66MB 下载

sklearn机器学习常用过程总结

由于前面对sklearn或多或少接触了一下, 但是不深入, 随着最近学习, 我下面介绍一下机器学习...

997

机器学习五大流派，九种常见算法优点场景举例

一、机器学习概览 1. 什么是机器学习？ 机器通过分析大量数据来进行学习。比如说, 不需要通...

1255

机器学习/数据挖掘/十大经典算法总结

一、C4.5决策树 1、背景知识介绍 分类树（决策树）是一种十分常用的机器学习的分类方法。他...

1386

机器学习实战-基本算法总结1

机器学习基本算法总结 ⇨ 监督学习——分类 代码在这, 基于python3 (原书代码是python2) 这...

2150

组织如何实现业务管理全程电子化？


移动办公专家“泛微”帮助您的组织实现业务管理全程电子化, 申请试用



机器学习笔记 监督学习算法小结(一)

前言坚持写博客半年多了, 但感觉自己没有写技术博客的天赋, 写出来的东西深度和广度都欠缺...

3883



机器学习算法一栏(模型、算法怎么选择)

1.引言 提起笔来写这篇博客, 突然有点愧疚和尴尬。愧疚的是, 工作杂事多, 加之懒癌严重, 导...

1223



《机器学习有意思！ 01》 - 世界上最简单的机器学习入门

本文首发于https://jizhi.im/blog/post/ml_is_fun_01 你是否也曾听人们谈起机器学习但是只有一个...

2.3万



机器学习简介及常用算法

概念：机器学习？机器学习是英文名称Machine Learning(简称ML)的直译。机器学习涉及概...
4

机器学习常见算法个人总结

写评论
朴素贝叶斯 参考[1] 事件A和B同时发生的概率为在A发生的情况下发生B或者在B发生的情况下发...

oa办公管理系统

百度 收藏

蓝桥杯常用算法汇总

&ar...emory.h>或<string.h>void *memset(void *s, int ch, size_t n); #include &...
微博

机器学习10大经典算法简介

C4.5 QQ 算法是机器学习算法中的一种分类决策树算法,其核心算法是ID3算法. C4.5算法继承了I...

机器学习入门----以阿里移动推荐算法大赛为例

从数据中学的模型的过程称为“学习”（learning）或“训练”（training），这个过程通过执行某个学...

机器学习综述

2017年04月18日 1022KB 下载

机器学习中特征选择概述

转载出处：http://blog.csdn.net/iezengli/article/details/32686803 一、背景 1) 问题 在机器学习的...

组织如何实现业务管理全程电子化？

移动办公专家“泛微”帮助您的组织实现业务管理全程电子化，申请试用

机器学习10大经典算法

1、C4.5 机器学习中，决策树是一个预测模型；他代表的是对象属性与对象值之间的一种映射关...

机器学习——01（sklearn工具箱机器学习算法的大致分类）

机器学习算法分类：（具体算法选择可根据样本量大小以及下面的特征进行选择） 1.无监督算法...

常用数据挖掘算法总结及Python实现 [强烈推荐]

2018年06月01日 31.36MB 下载

VB常用算法大全 收录了众多常用算法

2009年09月02日 3.83MB 下载

机器学习经典算法总结（3）——特征选择

一、特征的分类1. 相关特征：对当前学习任务有用的属性。2. 无关特征：对当前学习任务没有用...

招聘海归硕士

百度广告

- 简单易学的机器学习算法——极限学习机(ELM)

极限学习机(Extreme Learning Machine) ELM

4

8.6万
- 简单易懂的机器学习算法——K-Means算法

写评论

4.6万
- 一、K-Means算法的简介

目录

【机器学习理论】第1部分 机器学习基础

收藏

920

对于机器学习算法也使用了一段时间，现在进行一个系统的学习和总结，也给自己补缺补差，回...
- 25个Java机器学习工具包

微信

2129

本文介绍了25个Java机器学习工具&库： Weka集成了数据挖掘工作的机器学习算法。这些算...

微博
- 机器学习常用算法总结（二）

QQ

60

接上一篇总结，这次简单总结（朴素贝叶斯）Naive Bayes 在学习NB之前要先明白两个概念：先...
- 程序员不会英语怎么行？

老司机教你一个数学公式秒懂天下英语
- 机器学习常用算法总结（一）

73

最近一直在学习机器学习的算法知识，由于是转专业的找着方面的工作的，所以决定通过写博客...
- 机器学习之常用算法总结

426

原文来自： 原文链接：IT经理网 机器学习无疑是当前数据分析领域的一个热点内容。很多人在...
- Opencv3.0各种机器算法总结

401

1、正态贝叶斯：normal Bayesian classifier 2、K最近邻：k nearest neighbors classifier 3、支...
- 蓝桥杯C_++/Java程序设计常用算法&技巧总结

2082

蓝桥杯C_++/Java程序设计常用算法&技巧总结
- 十种深度学习算法要点及代码解析

3.3万

前言 谷歌董事长施密特曾说过：虽然谷歌的无人驾驶汽车和机器人受到了许多媒体关注，但是这...
- 没有更多推荐了, 返回首页

4

写评论

目录

收藏

微信

微博

QQ

勋章：



瓦努阿图移民



最新文章

监督学习框架通用处理流程

Label的设计以及特征列

深度学习资料整理

经典诗词汇总

大脑中的支持向量机

博主专栏

 C++知识总结

阅读量：1114741

 算法与数据结构

阅读量：1136251

个人分类

C++

算法与数据结构

工作相关

机器学习

深度学习

展开

归档

2018年1月

2017年12月

2017年6月

2016年12月

2016年11月

展开