

Super Resolution based Photo Enhancement by Google

2019年6月4日 18:32

<https://sites.google.com/view/handheld-super-res/>

拍摄时实际拍摄一系列图像，不需要理想化的情况即可进行工作，例如抖动，不好的光照。

可以获取更多图像信息并提取出来，拼接成高质量的图像。复杂的动作会导致不需要的重影，需要构建突出显示的鲁棒性掩码检测哪些区域会产生不需要的伪影，保证高质量的输出。

该方法不需要构建神经网络，只需要100毫秒，10张1200w像素图片/s。

这个项目也可以生成高质量的地面上实况数据，with很好的基线。

可口可乐 - 纪录片

2019年6月10日 15:24

日均8亿罐，浓缩液加水稀释，再加甜味剂与二氧化碳。

水中性化处理，完成标准化。甜味剂使用高果糖玉米糖浆，便宜。

10%含糖量，二氧化碳含量固定。

乌合之众

2019年6月11日 23:37

序

个体的聚集会产生新的心理特征且与种族特征不符，可能为模因。个体的意识行为可能会被群体的无意识行为取

如果要验证某种现象，就不应该考虑验证行为是否会伤害利益群体，归属于某个学派也可能导致偏见。

对于一个民族而言，最重要的就是对重大变革的狂热。WHY？制度与法律是人类外在性格的表达，所以无法改变

某些情况下，虚假比真实包含了更多的真相，就文明的演变而言，实践价值比理论价值更重要。

群体的无意识行为是神秘的，也是影响极大的。

引言

文明的变化是观念，理念和信念的变化。旧的宗教政治信仰，社会信念会被现代文明与科技改变，大众的声音进

群众 = 无意识残忍群体，只在毁灭阶段对文明起作用。

统治阶级了解群体特征，群体没有自己的意见，也不完全理性。

1.1 群体的一般特征 - 群体精神统一的心理定律

群体中个体的思想都会趋同，意识人格逐渐消失。群体特征有些与个体相似，有些独有的却相反，导致个体的感

变化：个体思考从多数量级出发，群体思想的传染导致为集体利益牺牲个人利益，暗示性破坏个体某些能力后大

代。

变人类的性格。此处存疑！

入政治。

受，思想，行为变化。个体的低层次无意识力量组成了群体特征。

幅度提升另一部分能力。成为群体后一个人的文明水平会大幅度下降。群体的优劣取决于个体组成群体后暴露出来的暗藏的性质

拉伸训练

2019年8月14日 18:56

#拉伸 一日三次

1. 颈部拉伸 手臂两侧拉伸接大字x轴左右转体 下拉韧带 50s一次组间10s
2. 手臂拉伸向上接钩住向左右 后拉韧带 大字z轴左右转体拉伸 50s一次组间10s
3. 倒立空中漫步前后探步 蟠体折叠，头不离地 剪刀腿两侧拉伸

#引体悬垂 脊柱周边肌肉拉伸 一周三次

悬垂 左右摇摆 前后摇摆 旋转

#颈部前伸

颈部弹力绳前后发力，左右拉伸

#肩膀

小臂外旋

外旋飞鸟

外旋划船

外旋飞鸟

#背

转肩

弓步上伸

远端拱桥

背两头起

骨骼，骨骼肌生长：

赖氨酸 亮氨酸 V-DA

Clean Code

2020年9月21日 14:50

Chapter 1 Clean Code

让营地比你来时更干净

Chapter 2 Naming

变量、函数或类的名称应该已经答复了所有的大问题。它该告诉你，它为什么会存在，它做什么事，应该怎么用。如果名称需要注释来补充，那就不算是名副其实。

程序员必须避免留下掩藏代码本意的错误线索。应当避免使用与本意相悖的词。

人类长于记忆和使用单词。大脑的相当一部分就是用来容纳和处理单词的。单词能读得出来。人类进化到大脑中有那么大的一块地方用来处理言语，若不善加利用，实在是种耻辱。

类名和对象名应该是名词或名词短语，如Customer、WikiPage、Account和AddressParser。避免使用Manager、Processor、Data或Info这样的类名。类名不应当是动词。

方法名应当是动词或动词短语，如postPayment、deletePage或save。

给每个抽象概念选一个词，并且一以贯之。

Chapter 3 Function

函数应该做一件事。做好这件事。只做这一件事。

要确保函数只做一件事，函数中的语句都要在同一抽象层级上。

我们想要让每个函数后面都跟着位于下一抽象层级的函数，这样一来，在查看函数列表时，就能循抽象层级向下阅读了。我把这叫做向下规则。

别害怕长名称。长而具有描述性的名称，要比短而令人费解的名称好。长而具有描述性的名称，要比描述性的长注释好。使用某种命名约定，让函数名称中的多个单词容易阅读，然后使用这些单词给函数取个能说清其功用的名称。

如果没有参数，就是小菜一碟。如果只有一个参数，也不太困难。有两个参数，问题就麻烦多了。如果参数多于两个，测试覆盖所有可能值的组合简直让人生畏。

输出参数比输入参数还要难以理解。读函数时，我们惯于认为信息通过参数输入函数，通过返回值从函数中输出。我们不太期望信息通过参数输出。

函数承诺只做一件事，但还是会做其他被藏起来的事。有时，它会对自己类中的变量做出未能预期的改动。有时，它会把变量搞成向函数传递的参数或是系统全局变量。无论哪种情况，都是具有破坏性的，会导致古怪的时序性耦合及顺序依赖。

重复可能是软件中一切邪恶的根源。

只要函数保持短小，偶尔出现的return、break或continue语句没有坏处，甚至还比单入单出原则更具有表达力。另外一方面，goto只在大函数中才有道理，所以应该尽量避免使用。

Chapter 4 Comment

带有少量注释的整洁而有表达力的代码，要比带有大量注释的零碎而复杂的代码像样得多。与其花时间编写解释你搞出的糟糕的代码的注释，不如花时间清洁那堆糟糕的代码。

Chapter 5 Format

好的软件系统是由一系列读起来不错的代码文件组成的。它们需要拥有一致和顺畅的风格。

Chapter 6 Object and Data Structure

Contextual Bandits: LinUCB

2020年10月19日 16:57

推荐系统选择商品展现给用户，并期待用户的正向反馈（点击、成交）。然而推荐系统并不能提前知道用户在观察到商品之后如何反馈，也就是不能提前获得本次推荐的收益，唯一能做的就是不停地尝试，并实时收集反馈以便更新自己试错的策略。目的是使得整个过程损失的收益最小。这一过程就类似与一个赌徒在赌场里玩老虎机赌博。赌徒要去摇老虎机，走进赌场一看，一排老虎机，外表一模一样，但是每个老虎机吐钱的概率可不一样，他不知道每个老虎机吐钱的概率分布是什么，那么每次该选择哪个老虎机可以做到最大化收益呢？这就是多臂赌博机问题（Multi-armed bandit problem, MAB）。

MAB问题的难点是Exploitation-Exploration(E&E)两难的问题：对已知的吐钱概率比较高的老虎机，应该更多的去尝试(exploitation)，以便获得一定的累计收益；对未知的或尝试次数较少的老虎机，还要分配一定的尝试机会（exploration），以免错失收益更高的选择，但同时较多的探索也意味着较高的风险（机会成本）。

Bandit算法是一类用来实现Exploitation-Exploration机制的策略。根据是否考虑上下文特征，Bandit算法分为context-free bandit和contextual bandit两大类。

在此，重点介绍一下UCB方法的基本思想。在统计学中，对于一个未知量的估计，总能找到一种量化其置信度的方法。最普遍的分布正态分布（或曰高斯分布）

，其中的

就是估计量的期望，而

则表示其不确定性（

越大则表示越不可信）。比如你掷一个标准的6面色子，它的平均值是3.5，而如果你只掷一次，比如说到2，那你对平均值的估计只能是2，但是这个置信度应该很低，我们可以知道，这个色子的预估平均值是2，而以95%的置信区间在[1.4,5.2]。

UCB (Upper Confidence Bound - 置信上限) 就是以收益（bonus）均值的置信区间上限代表对该arm未来收益的预估值：

- 对于未知或较少尝试的arm，尽管其均值可能很低，但是由于其不确定性会导致置信区间的上界较大，从而有较大的概率触发exploration
- 对于已经很熟悉的arm(尝试过较多次)，更多的是触发exploitation机制：如果其均值很高，会获得更多的利用机会；反之，则会减少对其尝试的机会

From <<https://zhuanlan.zhihu.com/p/35753281>>

在推荐系统中，通常把待推荐的商品作为MAB问题的arm。UCB这样的context-free类算法，没有充分利用推荐场景的上下文信息，为所有用户的选择展现商品的策略都是相同的，忽略了用户

作为一个个活生生的个体本身的兴趣点、偏好、购买力等因素都是不同的，因而，同一个商品在不同的用户、不同的情景下接受程度是不同的。故在实际的推荐系统中，context-free的MAB算法基本都不会被采用。

与context-free MAB算法对应的是Contextual Bandit算法，顾名思义，这类算法在实现E&E时考虑了上下文信息，因而更加适合实际的个性化推荐场景。

- 每个arm学习一个独立的模型（context只需要包含user-side和user-arm interaction的特征，不需要包含arm-side特征）；而传统在线学习为整个业务场景学习一个统一的模型
- 传统的在线学习采用贪心策略，尽最大可能利用已学到的知识，没有exploration机制（贪心策略通常情况下都不是最优的）；LinUCB则有较完善的E&E机制，关注长期整体收益

From <<https://zhuanlan.zhihu.com/p/35753281>>

蒙特卡洛梯度估计方法 (MCGE)

2020年10月19日 17:23

MCMC采样算法

对于给定的概率分布 $p(x)p(x)$ ，我们希望能有便捷的方式生成它对应的样本。由于马尔科夫链能收敛到平稳分布，于是一个很的漂亮想法是：

如果我们能构造一个转移矩阵为 P 的马尔科夫链，使得该马尔科夫链的平稳分布恰好是 $p(x)p(x)$ ，那么我们从任何一个初始状态 x_0 出发沿着马尔科夫链转移，得到一个转移序列 $x_0, x_1, x_2, \dots, x_n, x_{n+1}, \dots$ ，如果马尔科夫在第 n 步已经收敛了，于是我们就得到了 $p(x)p(x)$ 的样本 $x_n, x_{n+1}, \dots, x_n, x_{n+1}, \dots$ 。

From <<https://blog.csdn.net/qy20115549/article/details/79635621>>

如果非周期马尔科夫链的转移矩阵 P 和分布 $\pi(x)\pi(x)$ 满足

$$\pi(i)P_{ij} = \pi(j)P_{ji} \text{ for all } i, j$$

则 $\pi(x)\pi(x)$ 是马尔科夫链的平稳分布，上式被称为细致平稳条件(detailed balance condition)。其实这个定理是显而易见的，因为细致平稳条件的物理含义就是对于任何两个状态 i, j ,

从 i 转移出去到 j 而丢失的概率质量，恰好会被从 j 转移回 i 的概率质量补充回来，所以状态 i 上的概率质量 $\pi(i)\pi(i)$ 是稳定的，从而 $\pi(x)\pi(x)$ 是马尔科夫链的平稳分布。

朴素贝叶斯推断 Bayesian inference

2020年10月19日 17:39

$$P(B_i | A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^n P(B_j)P(A|B_j)}$$

Baidu 百度

求A中Bi

Bi先验概率 * Bi中A比例

/

不Bi先验概率 * 不Bi中A比例

一些文献中把 $P(B[1])$ 、 $P(B[2])$ 称为基础概率， $P(A | B[1])$ 为击中率， $P(A | B[2])$ 为误报率[1]

From <<https://baike.baidu.com/item/%E8%B4%9D%E5%8F%B6%E6%96%AF%E5%AE%9A%E7%90%86>>

Naïve Bayes 算法，又叫朴素贝叶斯算法，朴素：特征条件独立；贝叶斯：基于贝叶斯定理。所谓朴素，就是整个形式化过程只做最原始假设

假设某个体有n项特征 (Feature)，分别为 F_1, F_2, \dots, F_n 。现有m个类别 (Category)，分别为 C_1, C_2, \dots, C_m 。贝叶斯分类器就是计算出概率最大的那个分类，也就是求下面这个算式的最大值：

$$P(C|F_1F_2\dots F_n) = P(F_1F_2\dots F_n|C)P(C) / P(F_1F_2\dots F_n)$$

由于 $P(F_1F_2\dots F_n)$ 对于所有的类别都是相同的，可以省略，问题就变成了求 $P(F_1F_2\dots F_n|C)P(C)$ 的最大值。

朴素贝叶斯分类器则是更进一步，假设所有特征都彼此独立，因此

$$P(F_1F_2\dots F_n | C)P(C) = P(F_1 | C)P(F_2 | C) \dots P(F_n | C)P(C)$$

4. 拉普拉斯平滑 (Laplace smoothing)

也就是参数为1时的贝叶斯估计，当某个分量在总样本某个分类中（观察样本库/训练集）从没出现过，会导致整个实例的计算结果为0。为了解决这个问题，使用拉普拉斯平滑/加1平滑进行处理。

它的思想非常简单，就是对先验概率的分子（划分的计数）加1，分母加上类别数；对条件概率分子加1，分母加上对应特征的可能取值数量。这样在解决零概率问题的同时，也保证了概率和依然为1。

Thompson sampling 汤普森采样

2020年10月19日 22:46

汤普森抽样 (Thompson Sampling) 的优势在于，随着我们获得越来越多的信息，搜索量会趋于减少，这模仿了我们希望在较少的搜索中获得尽可能多的信息的问题之间的取舍。因此，当我们拥有更少的数据时，该算法趋向于更加“面向搜索”，而当我们拥有大量数据时，该算法具有“面向探索”的趋势。

From <<https://www.geeksforgeeks.org/introduction-to-thompson-sampling-reinforcement-learning/>>

每次对每个类随机抽取一个样本，推荐 θ 最大的
获取reward
根据reward更新 θ 分布

A Tutorial on Thompson Sampling

Beta分布有很多与生俱来的性质，随着我们观测到的结果越多，分布的置信区间就越窄，比方说最开始 $\alpha=1, \beta=1$ ，当我们观测到3次正回报，5次负回报后， $(\alpha, \beta)=(4, 6)$ ，就产生了上图中action 3的概率分布。当我们观测到599次回报和399次零回报的动作后， $(\alpha, \beta)=(600, 400)$ ，就产生了上图中action 1的概率分布。

我们可以在 $K=10$ 的多臂赌博机上对比一下Epsilon-Greedy和Thompson Sampling的实验效果，当轮数较少时，Thompson Sampling可能会以比较大的概率进行探索，较少的情况使用最佳动作，产生比较高的regret，但是轮数上升后，Thompson Sampling的尝试收敛了，总regret数也随着收敛。

From <<https://cloud.tencent.com/developer/news/100783>>

Weight Uncertainty in Neural Networks

2020年10月19日 17:18

我们神经网络中的所有权重都由可能值上的概率分布表示，而不是像规范那样具有单个固定值

岭回归(Ridge Regression)

2020年10月19日 17:06

- 模型的解释能力：如果模型中的特征之间有相互关系，这样会增加模型的复杂程度，并且对整个模型的解释能力并没有提高，这时，我们就要进行特征选择。

以上的这些问题，主要就是表现在模型的方差和偏差问题上，这样的关系可以通过下图说明

$$\text{线性回归的损失函数: } J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2$$

$$\text{岭回归的损失函数: } J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda \sum_{j=1}^n \theta_j^2$$

$$\text{Lasso回归的损失函数: } J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda \sum_{j=1}^n |\theta_j|$$

其中 λ 称为正则化参数，如果 λ 选取过大，会把所有参数 θ 均最小化，造成欠拟合，如果 λ 选取过小，会导致对过拟合问题解决不当，因此 λ 的选取是一个技术活。

岭回归与Lasso回归最大的区别在于岭回归引入的是L2范数惩罚项，Lasso回归引入的是L1范数惩罚项，Lasso回归能够使得损失函数中的许多 θ 均变成0，这点要优于岭回归，因为岭回归是要所有的 θ 均存在的，这样计算量Lasso回归将远远小于岭回归。

Boyer-Moore 投票算法

2020年10月3日 22:42

计数=0

比较值 = 第一位

计数+=1

相等计数+=1

不相等计数-=1

计数==0时换比较值=当前值

计数+=1

O(n)算法

证明：

当前一段计数==0时，出现最多次数（一半）

的值占了一半，其他潜在的正确答案必不可超过这个一半，如果答案是潜在的，后续必定过半，如果正确答案就是它，后续还是至少占一半。

class Solution:

```
def majorityElement(self, nums: List[int]) -> int:  
    if len(nums) == 1:  
        return nums[0]  
    tmp = nums[0]  
    cnt = 1  
    for i in range(1, len(nums)):  
        if cnt == 0:  
            tmp = nums[i]  
        if tmp == nums[i]:  
            cnt += 1  
        else:  
            cnt -= 1  
    return tmp
```

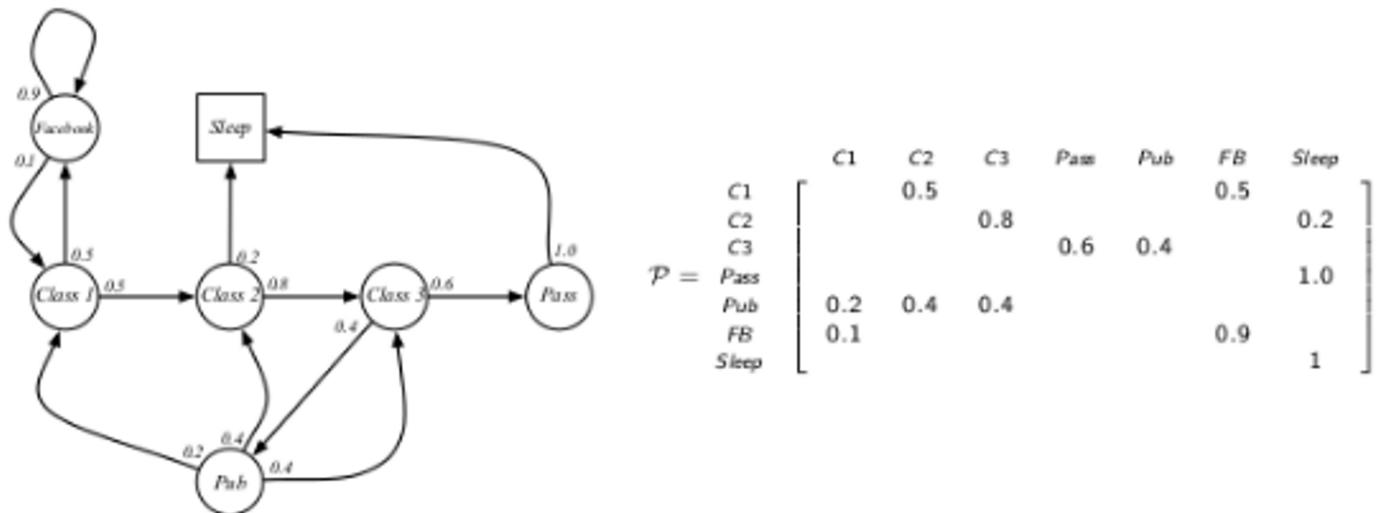
马尔科夫决策过程

2020年10月16日 2:20

在强化学习中，马尔科夫决策过程（Markov decision process, MDP）是对完全可观测的环境进行描述的，也就是说观测到的状态内容完整地决定了决策的需要的特征。几乎所有的强化学习问题都可以转化为MDP。

某一状态信息包含了所有相关的历史，只要当前状态可知，所有的历史信息都不再需要，当前状态就可以决定未来，则认为该状态具有马尔科夫性。

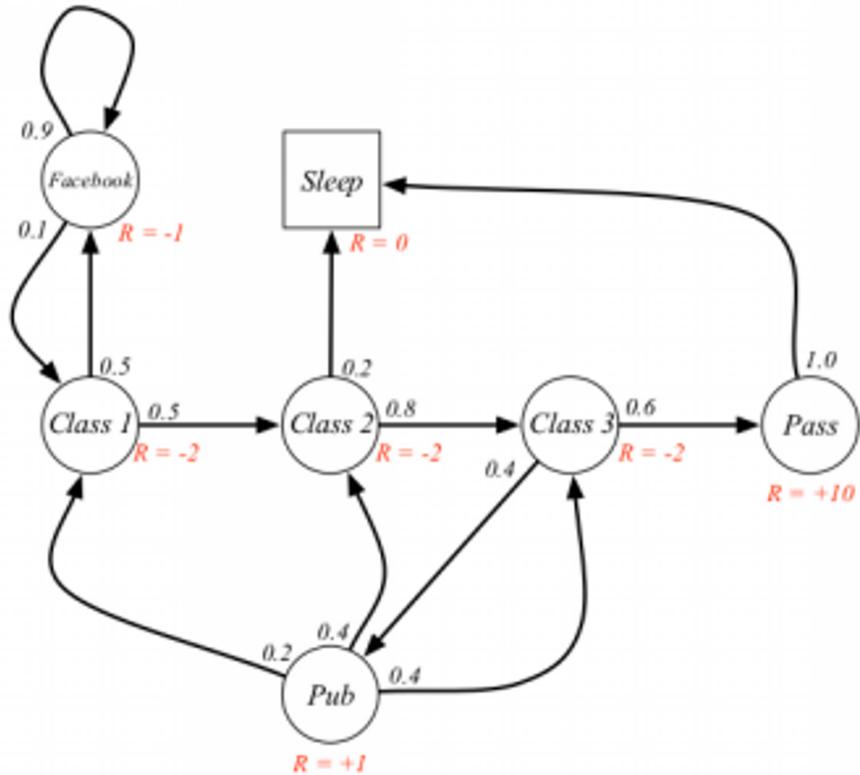
马尔科夫过程又叫马尔科夫链(Markov Chain)，它是一个无记忆的随机过程，可以用一个元组 $\langle S, P \rangle$ 表示，其中S是有限数量的状态集，P是状态转移概率矩阵。



马尔科夫奖励过程在马尔科夫过程的基础上增加了奖励R和衰减系数 γ : $\langle S, P, R, \gamma \rangle$ 。R是一个奖励函数。S状态下的奖励是某一时刻(t)处在状态s下在下一个时刻(t+1)能获得的奖励期望：

$$R_s = E[R_{t+1} | S_t = s]$$

衰减系数 Discount Factor: $\gamma \in [0, 1]$ ，它的引入有很多理由，David也列举了不少原因来解释为什么引入衰减系数，其中有数学表达的方便，避免陷入无限循环，远期利益具有一定的不确定性，符合人类对于眼前利益的追求，符合金融学上获得的利益能够产生新的利益因而更有价值等等。



收获 Return

定义：收获 G_t 为在一个马尔科夫奖励链上从 t 时刻开始往后所有的奖励的有衰减的总和。也有翻译成“收益”或“回报”。公式如下：

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

其中衰减系数体现了未来的奖励在当前时刻的价值比例，在 $k+1$ 时刻获得的奖励 R 在 t 时刻的体现出的价值是 $\gamma^k * R$ ， γ 接近 0，则表明趋向于“近视”性评估； γ 接近 1 则表明偏重考虑远期的利益。

价值函数给出了某一状态或某一行为的长期价值。

定义：一个马尔科夫奖励过程中某一状态的价值函数为从该状态开始的马尔可夫链收获的期望：

$$v(s) = E[G_t | S_t = s]$$

<i>States</i>	<i>C₁</i>	<i>C₂</i>	<i>C₃</i>	<i>Pass</i>	<i>Pub</i>	<i>FB</i>	<i>Sleep</i>
<i>Rewards</i>	-2	-2	-2	10	1	-1	0
<i>C₁</i>		0.5				0.5	
<i>C₂</i>			0.8				0.2
<i>C₃</i>				0.6	0.4		
<i>Pass</i>							1
<i>Pub</i>	0.2	0.4	0.4				
<i>FB</i>	0.1					0.9	
<i>Sleep</i>							1

<i>C₁ C₂ C₃ Pass Sleep</i>	$G_1 = -2 + (-2) * 1/2 + (-2) * 1/4 + 10 * 1/8 + 0 * 1/16 = -2.25$
<i>C₁ FB FB C₁ C₂ Sleep</i>	$G_1 = -2 + (-1) * 1/2 + (-1) * 1/4 + (-2) * 1/8 + (-2) * 1/16 + 0 * 1/32 = -3.125$
<i>C₁ C₂ C₃ Pub C₂ C₃ Pass Sleep</i>	$G_1 = -2 + (-2) * 1/2 + (-2) * 1/4 + (1) * 1/8 + (-2) * 1/16 + \dots = -3.41$
<i>C₁ FB FB C₁ C₂ C₃ Pub C₁ FB FB C₁ C₂ C₃ Pub C₂ Sleep</i>	$G_1 = -2 + (-1) * 1/2 + (-1) * 1/4 + (-2) * 1/8 + (-2) * 1/16 + (-2) * 1/32 + \dots = -3.20$

各状态价值的确定是很重要的，RL的许多问题可以归结为求状态的价值问题。因此如何求解各状态的价值，也就是寻找一个价值函数（从状态到价值的映射）就变得很重要了。

$$\begin{aligned}
v(s) &= \mathbb{E}[G_t \mid S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots \mid S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} + \dots) \mid S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma G_{t+1} \mid S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma v(S_{t+1}) \mid S_t = s]
\end{aligned}$$

$$v(s) = \mathbb{E}[R_{t+1} + \gamma v(S_{t+1}) \mid S_t = s]$$

通过方程可以看出由两部分组成，一是该状态的即时奖励期望，即时奖励期望等于即时奖励，因为根据即时奖励的定义，它与下一个状态无关；另一个是下一时刻状态的价值期望，可以根据下一时刻状态的概率分布得到其期望。如果用 s' 表示 s 状态下一时刻任一可能的状态，那么Bellman方程可以写成：

$$v(s) = \mathcal{R}_s + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'} v(s')$$

<https://zhuanlan.zhihu.com/p/28084942>

极大似然估计 + 最大后验概率

2020年10月19日 23:06

使用极大似然估计方法的两个条件

1. 我们假定数据服从某种已知的特定数据分布型。

2. 我们已经得到了一定的数据集。

From <<https://www.jianshu.com/p/e0eb4f4ccf3e>>

一个随机试验如有若干个可能的结果A, B, C, …, 若在一次试验中, 结果A出现了, 那么可以认为实验条件对A的出现有利, 也即出现的概率P(A)较大。极大似然原理的直观想法我们用下面例子说明。设甲箱中有99个白球, 1个黑球; 乙箱中有1个白球, 99个黑球。现随机取出一箱, 再从抽取的一箱中随机取出一球, 结果是黑球, 这一黑球从乙箱抽取的概率比从甲箱抽取的概率大得多, 这时我们自然更多地相信这个黑球是取自乙箱的。一般说来, 事件A发生的概率与某一未知参数

有关,

取值不同, 则事件A发生的概率

也不同, 当我们在一次试验中事件A发生了, 则认为此时的
值应是t的一切可能取值中使

达到最大的那一个, 极大似然估计法就是要选取这样的t值作为参数t的估计值, 使所选取的样本
在被选的总体中出现的可能性为最大。

From <<https://baike.baidu.com/item/%E6%9E%81%E5%A4%A7%E4%BC%BC%E7%84%B6%E4%BC%B0%E8%AE%A1>>

最大后验概率估计是最大似然和贝叶斯估计的结合

这里需要说明, 虽然从公式上来看 $M A P = M L E * P(\theta)$ $MAP=MLE*P(\theta)$, 但是这两种算法有本质的区别, 极大似然估计(MLE)将 θ 视为一个确定未知的值, 而最大后验概率估计(MAP)则将 θ 视为一个随机变量。

版权声明: 本文为CSDN博主「gcheney」的原创文章, 遵循CC 4.0 BY-SA版权协议, 转载请附上原
文出处链接及本声明。

原文链接: <https://blog.csdn.net/gcheney/article/details/108442861>

L_p 正则化

2021年5月25日 15:53

总结L1正则化和L2正则化：

L1范数：为x向量各个元素绝对值之和。

L2范数：为x向量各个元素平方和的1/2次方， L2范数又称Euclidean范数或Frobenius范数

L_p范数：为x向量各个元素绝对值_p次方和的1/_p次方。

From <<https://zhuanlan.zhihu.com/p/352437358>>

线形回归的L1正则化通常称为Lasso回归

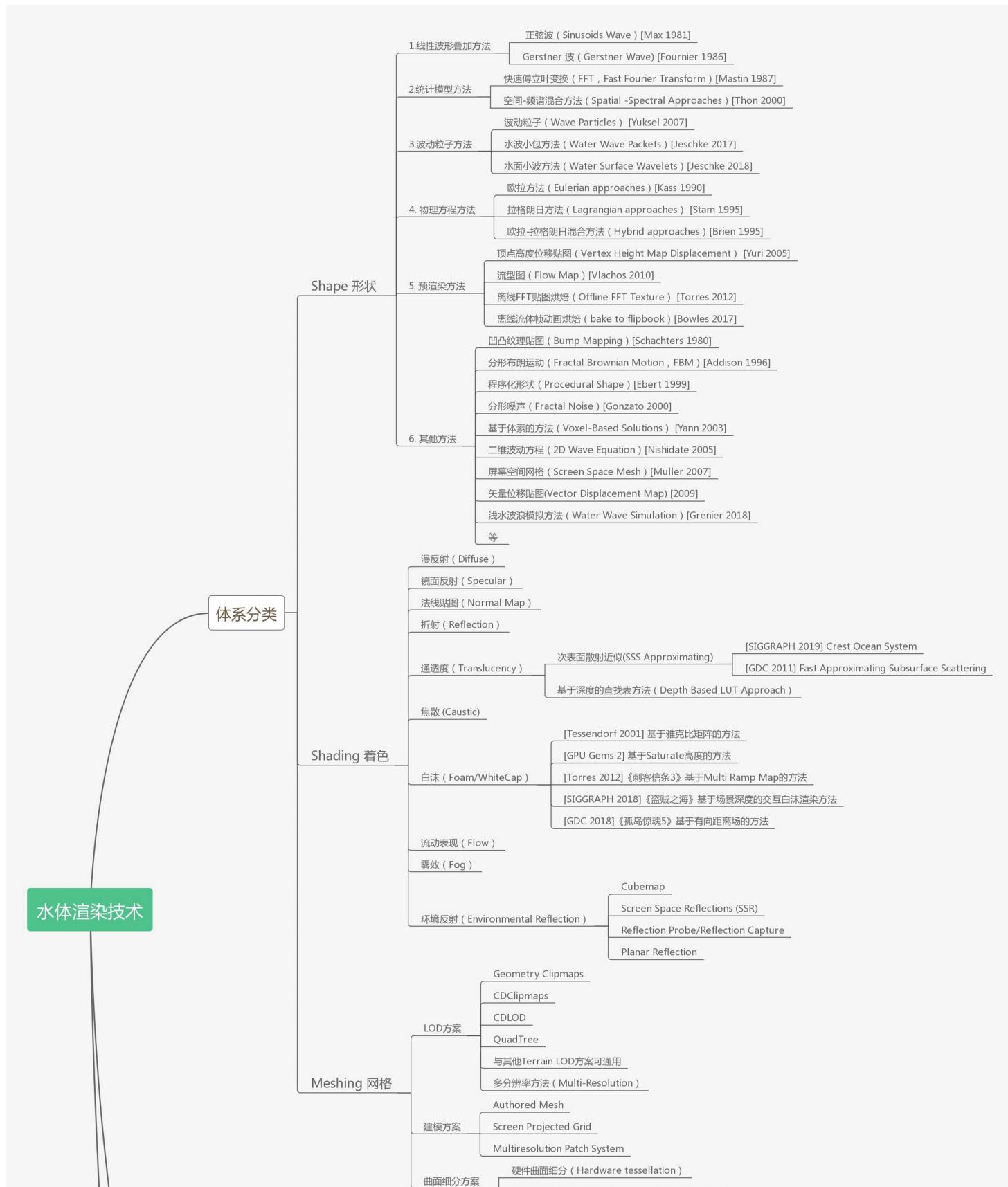
From <<https://zhuanlan.zhihu.com/p/352437358>>

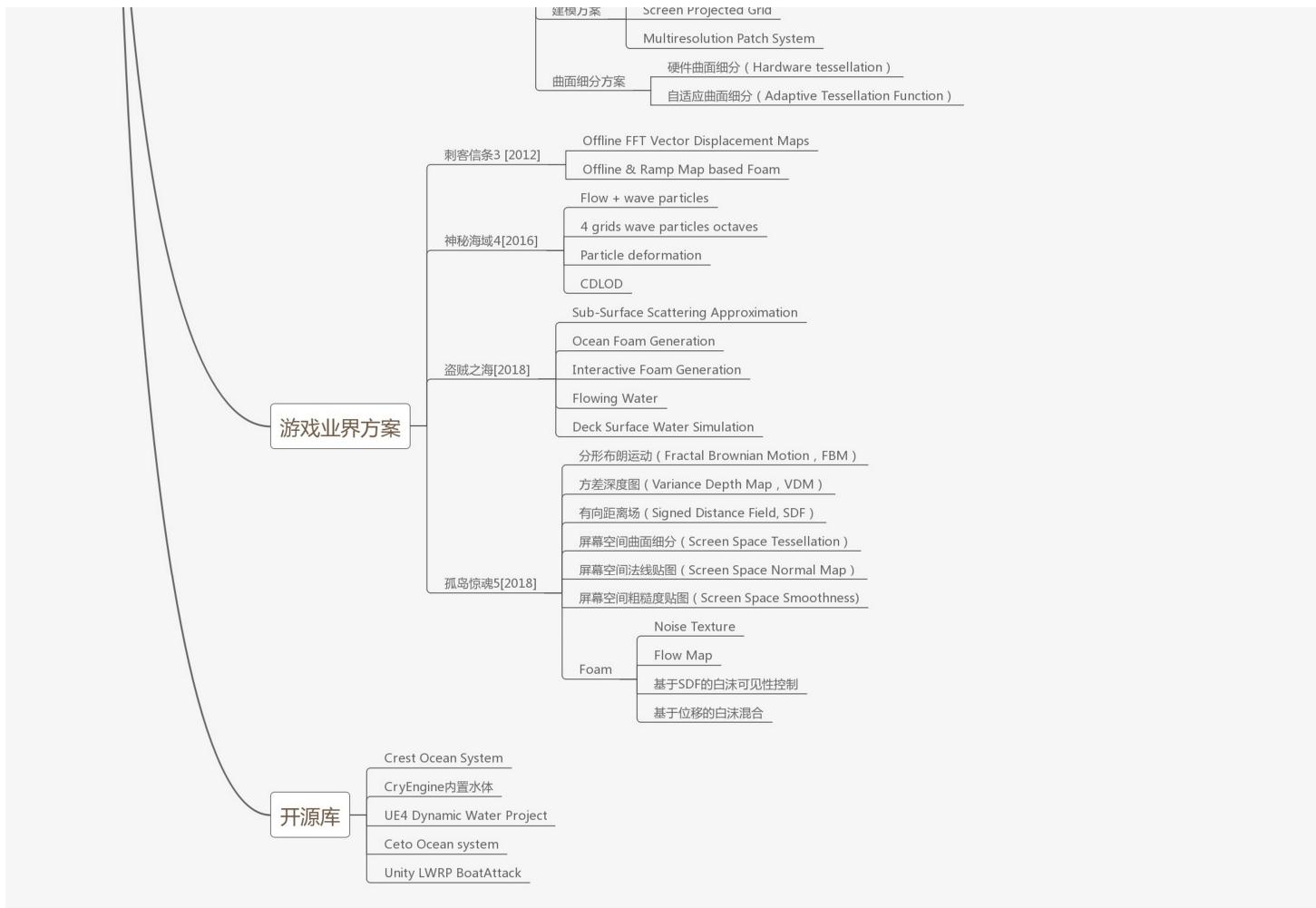
线形回归的L2正则化通常称为Ridge回归

From <<https://zhuanlan.zhihu.com/p/352437358>>

水体渲染

2021年1月21日 15:39





方法1：

线性波形函数叠加在水体每个点上，例如正弦/Gerstner

简单

方法2：

快速傅里叶变换，又称基于频谱方法，构造表面高度

使用理论或统计获得的海浪频谱，结合大量正弦波生成波形分布，逆FFT转到空间域，生成位移& 法线贴图

可离线/实时渲染，计算量较大，电影中使用离线 (before 2000)

方法3：

波动粒子 (Wave Particle) 方法最初由Yuksel于2007年[Yuksel 2007]提出，该方法的核心思想是采用粒子代表每一个水波，并允许波反射以及与动态对象的相互作用。

随后的[SIGGRAPH 2018]，Jeschke等人对Water Wave Packets进行了改进，提出了新的水面小波方法 (Water Surface Wavelets)。水面小波方法 (Water Surface Wavelets) 基于欧拉方法，自由度与空间区域有关，与波动本身无关。

方法，自由度与空间区域有关，与流动本身无关。

物理方法无法实时，只可用于电影的离线渲染，包括欧拉方法和拉格朗日方法

方法4：

预渲染方法

流形图，本质上是一种基于矢量场平移法线贴图的着色技术，或者可以理解为一种UV动画，核心思想是预烘焙2D方向信息到纹理，以在运行时基于UV采样，对流动感进行模拟。

<https://zhuanlan.zhihu.com/p/95917609>

图像模糊算法

2021年2月23日 12:43

模糊算法 Blur Algorithm	模糊品质 Quality	稳定性 Stability	性能 Performance	备注 Note
高斯模糊 Gaussian Blur	高	好	一般	基于卷积的原算法需要大量采样次数，即便做了线性 Separate，一次模糊迭代仍需要两个 pass。
方框模糊 Box Blur	高	好	一般	单次模糊计算可控，但能提供的模糊感不强，对于高品质的模糊表现则需要多次迭代的支撑。
Kawase 模糊 Kawase Blur	高	好	较好	渐变式的多 pass 模糊 kernel 设计，单次迭代即可带来一定的模糊质感，整体性价比较高。
双重模糊 Dual Blur	高	好	好	升降采样+渐变式的多 pass 模糊 kernel 设计，算法整体性价高，且优于标准 Kawase Blur。
散景模糊 Bokeh Blur	高	好	差	基于 Golden Angle 思路的实现，高品质的散景模糊需要大量迭代次数的支撑。
移轴模糊 Tilt Shift Blur	高	好	差	又称作镜头模糊。基于 Golden Angle 思路的实现，高品质的移轴模糊需要大量迭代次数的支撑。
光圈模糊 Iris Blur	高	好	差	基于 Golden Angle 思路的实现，高品质的镜头模糊需要大量迭代次数的支撑。
粒状模糊 Grainy Blur	一般	好	好	基于单次采样+随机 uv 进行粗粒度模糊的模拟，且在单 pass 下即可有合适的模糊表现，性能出色。
径向模糊 Radial Blur	高	好	一般	常用于速度感的表达，高品质的径向模糊需要多次迭代的支撑。
方向模糊 Directional Blur	高	好	一般	常用于方向感的表达，高品质的方向模糊同样需要多次迭代的支撑。

Ten Simple Rules to becoming a principal investigator

2021年2月4日 20:17

1. 创造力 - 思考新的想法，随时随地。捕获并不新颖或者错误的。让他们有时间发展。在形成雏形后再与他人交流。
2. 保证出版，一作优先。
3. 承担多方面的责任。
4. 发展额外技能，写作，合作和管理。
5. 专注于关键要素，评估单位时间收益。学会说不。
6. 接受失败，从中学习。
7. 发展个人品牌。
8. 相信自己，确保个人和团队的身心健康。
9. 建立人际网络。
10. 知道如何退出。

TCP 拥塞控制

2021年2月6日 17:27

tcp持续发送报文，到达预定时间未发到则重发，中间的尺寸为拥塞窗口。

当丢包率高时判断发生拥塞，减小窗口。

一般来说，窗口控制方法为乘法减小，加法增大。

如何提问

2021年2月13日 15:50

提问之前：

- 搜索旧问答
- 文档阅读 FAQ
- 阅读错误日志 源码

提问时：

- 谨慎选择提问场合
- 主题一致
- 难度水准一致
- 避免私信
- Google - Stack Overflow
- 使用有意义的明确标题
- 清晰，正确，精准，无错误的语句
English is not my native language, please excuse typing errors.
- 使用标准的文本格式描述问题
- 精确描述问题
 - 描述目标而非直接描述问题本身
 - 症状
 - 环境，如何重现
 - 如何研究，理解，处理
 - 最近可能导致该问题的变更
 - 不要说找到bug/盲目猜测
- 保持礼貌与感谢
- 在问题解决后补充说明

解读回复：

- RTFM - Read the fucking manual
- STFW - Search the fucking web
- 看不懂回复时先尝试搞懂

如何回答：

- 礼貌善良
- 不开玩笑
- 询问关键细节

保证回复质量，不重复
展现过程而非结果

Automated ML

2021年2月14日 12:24

Forward

AutoML 的目的在于节约时间

Full automation为研究动机和长期工程目标

目前目标是将大部分工作semi-auto来减少人类在循环中的介入

手段是开发强大的工具，使ML过程更加系统化且高效

目前AutoML可以在一些目标中表现优于人类专家，这表明一个有前景的开端

该书会包括：超参数优化， meta-learning， 神经结构检索

Preface

AutoML目的在于使网络结构/训练策略/正则化方法/超参数这些 选择都可以在数据驱动的基础上
自动化，标准化进行。降低使用ML技术的人力成本和技术门槛。

ADF Test

2021年3月9日 14:01

adf检验是用来检验序列是否平稳的方式，一般来说是时间序列中的一种检验方法。
python中可使用现成的工具statsmodels来实现adf检验。

```
From <https://www.cnblogs.com/wynlfd/p/8862482.html>
import numpy as np
import statsmodels.tsa.stattools as ts
x = np.array([1, 2, 3, 4, 5, 6, 7])
result = ts.adfuller(x, 1)
print result
```

From <<https://www.cnblogs.com/wynlfd/p/8862482.html>>

民间科学爱好者的行为及心理分析

2021年3月15日 11:32

Def

所谓民间科学爱好者，是指在科学共同体之外进行所谓科学研究的一个特殊人群，他们或者希望一举解决某个重大的科学问题，或者试图推翻某个著名的科学理论，或者致力于建立某种庞大的理论体系，但是他们却不接受也不了解科学共同体的基本范式，与科学共同体不能达成基本的交流。总的来说，他们的工作不具备科学意义上的价值。

这个定义中，只有“不能交流”可以算做可观察的外部行为，而“希望”、“试图”和“致力于”等都与动机、目的等心理因素有关，不能直接观察。对此，只有通过对其行为的分析，和对它本人的访谈予以判断。可以想象，研究民间科学爱好者有如下几种材料：1，与民科的直接往来；2，对其本人及相关人物的访谈；3，大众媒体的报道；4，民科的文本（“科研论文”）。实际上，当我们以科学家为对象进行社会学研究时，同样要借助于这几种材料。

在定义了民科之后，笔者还定义了业余科学爱好者（science amateur）。这一群体同样人数众多，同样在科学共同体之外从事科学活动，比如天文爱好者，他们经常观测星象，也能发现一些新的天体如小行星、彗星；或者如生物爱好者，他们制作标本，也能发现一些新的生物种类。他们与民科的区别在于：从外部看，他们能够与科学共同体进行交流；从心理上分析，他们不想推翻现有的科学体系，一鸣惊人，而只是出于爱好，做一些具体的科学工作。民间科学爱好者、业余科学爱好者乃至科学共同体内部的职业科学家是一个连续谱，其间不存在截然分明的界限。

From <<http://shc2000.sjtu.edu.cn/0403/minjian.htm>>

核心心理特征：偏执->交流缺失/不接受反面意见+回避质疑+基于误解夸大别人错误/自大/牺牲+自我感动/行为艺术模仿学术行为+巫术带来的替代性满足

ARIMA

2021年3月17日 13:46

Auto regressive integrated moving average

基于时间序列历史与其上预测误差进行预测

用于非季节性时间序列预测

Hyper-Parameter: p d q

p: 自回归阶数-当前数值与前p各历史数值相关

d: 最小差分阶数-一般为1, 使序列平稳但丢失自相关性

q: 滑动平均项阶数-当前数值与前q个历史预测误差相关

$$Y_t = c + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

即: 被预测变量 Y_t = 常数+ Y 的p阶滞后的线性组合 + 预测误差的q阶滞后的线性组合

From <<https://cloud.tencent.com/developer/article/1646121>>

Kurtosis Cal Code

2021年3月29日 15:19

%<https://blog.csdn.net/u013649834/article/details/51527552> for kurtosis

知识图谱

2021年4月26日 14:46

知识图谱是由Google公司在2012年提出来的一个新的概念。从学术的角度，我们可以对知识图谱给一个这样的定义：“知识图谱本质上是语义网络（Semantic Network）的知识库”。但这有点抽象，所以换个角度，从实际应用的角度出发其实可以简单地把知识图谱理解成多关系图（Multi-relational Graph）。

From <<https://zhuanlan.zhihu.com/p/80280367>>

信息抽取的难点在于处理非结构化数据。在下面的图中，我们给出了一个实例。左边是一段非结构化的英文文本，右边是从这些文本中抽取出来的实体和关系。在构建类似的图谱过程当中，主要涉及以下几个方面的自然语言处理技术：

- a. 实体命名识别（Name Entity Recognition）
- b. 关系抽取（Relation Extraction）
- c. 实体统一（Entity Resolution）->合并实体 稀疏化图谱
- d. 指代消解（Coreference Resolution）代词->实体

From <<https://zhuanlan.zhihu.com/p/80280367>>

1. 需要哪些实体、关系和属性？ 2. 哪些属性可以做为实体，哪些实体可以作为属性？ 3. 哪些信息不需要放在知识图谱中？

From <<https://zhuanlan.zhihu.com/p/80280367>>

Retinex

2021年5月10日 17:09

Retinex是一种常用的建立在科学实验和科学分析基础上的图像增强方法，它是Edwin.H.Land于1963年提出的。就跟Matlab是由Matrix和Laboratory合成的一样，Retinex也是由两个单词合成的一个词语，他们分别是retina 和cortex，即：视网膜和皮层。Land的retinex模式是建立在以下三个假设之上的：

真实世界是无颜色的，我们所感知的颜色是光与物质的相互作用的结果。我们见到的水是无色的，但是水膜—肥皂膜却是显现五彩缤纷，那是薄膜表面光干涉的结果。

每一颜色区域由给定波长的红、绿、蓝三原色构成的；

三原色决定了每个单位区域的颜色。

Retinex理论的基础理论是物体的颜色是由物体对长波（红色）、中波（绿色）、短波（蓝色）光线的反射能力来决定的，而不是由反射光强度的绝对值来决定的，物体的色彩不受光照非均匀性的影响，具有一致性，即retinex是以色感一致性（颜色恒常性）为基础的。不同于传统的线性、非线性的只能增强图像某一类特征的方法，Retinex可以在动态范围压缩、边缘增强和颜色恒常三个方面达到平衡，因此可以对各种不同类型的图像进行自适应的增强。

40多年来，研究人员模仿人类视觉系统发展了Retinex算法，从单尺度Retinex算法，改进成多尺度加权平均的MSR算法，再发展成彩色恢复多尺度MSRCR算法

版权声明：本文为CSDN博主「大熊背」的原创文章，遵循CC 4.0 BY-SA版权协议，转载请附上原文出处链接及本声明。

原文链接：<https://blog.csdn.net/lz0499/article/details/81154937>

Tensor Product

2021年5月12日 15:44

Weighted Automata

2021年5月12日 20:04

加权有限自动机为每个过渡添加权重（即 δ 是映射 $Q \times \Sigma \times Q \rightarrow R$ ）。通过加权FA的接受路径的权重是沿路径的过渡权重的乘积。加权FA定义了加权语言或字符串上的分布，其中字符串的权重是字符串的所有接受路径的权重的总和。

在概率FA中，每个状态都具有所有现有转换的权重之和为1的属性。此外，还有一个特殊的停止符号，我们写为 $\langle / s \rangle$ ，我们假设它出现在每个字符串的末尾，并且仅出现在每个字符串的末尾。然后，加权FA还定义了字符串上的概率分布。