

提示：采用手机 safari 微软翻译技术

## 1. 在线平台信息评级系统的设计：来自两个实验的证据

作者: [nikhil garg](#), [ramesh johari](#)

**摘要:** 平台严重依赖评级系统来了解市场参与者的质量。然而，在实践中，这些评级往往高度膨胀，大幅减少了可用于区分质量的信号。我们考虑两个问题：第一，评级制度能否通过改变评级制度中等级的含义和相对重要性，更好地区分质量？其次，如果是这样，平台在评级系统的设计中应该如何优化这些选择？我们首先分析了在网上劳动力市场上进行的随机对照试验的结果，在这个试验中，反馈表格中增加了一个额外的问题。在治疗条件之间，我们会改变问题的措辞和答案的选择。我们进一步在具有类似结构的亚马逊机械土耳其人身上进行了一次实验，以证实劳动力市场的发现。我们的测试显示，目前的通胀规范其实可以通过重新锚定评级制度水平的含义来对抗。特别是，正倾斜的比例，并提供了具体的解释，每个标签意味着收益率评级分布，提供了更多关于质量的信息。其次，我们开发了一个理论框架，通过选择答案标签及其解释，最大限度地提高到真正的底层质量分布的收敛速度，来优化评级系统的设计。最后，我们使用经验校准模型进行模拟，并利用这些模型研究对优化评级系统设计的影响。我们的模拟表

明, 我们的建模和优化方法可以大大提高通过基线设计获得的信息的质量。总体而言, 我们的研究表明, 可以设计在实践中提供信息的评级系统, 并演示如何以原则性的方式设计这些评级系统。

少

2018 年 10 月 30 日提交;最初宣布 2018 年 10 月。

## 2. 在 atlas 中为云联合分布式存储

**作者 :**frank berghaus, kevin casteels, alessandro di girolamo, colson driemel, marcus ebert, fabrizio furano, fernando galindo, mario lassnig, colin leavett-brown, michael paterson, cedric serfon, rof seuster, randall sobie, reda tafirout, ryan paul taylor

**摘要:** 在云计算中心运行的应用程序的输入数据可以存储在遥远的存储库中, 通常会将热门数据的多个副本存储在多个站点上。查找和检索远程数据可能具有挑战性, 我们相信联合存储可以解决此问题。联盟将根据 geoip 信息查找数据的最接近的副本。目前, 我们正在使用由 cern it 开发的软件解决方案——动态数据联合 dynafed。dynafed 支持多个连接协议的行业标准, 如 **amazon 的 s3**、microsoft 的 azure 以及 webdav 和 http。dynafed 作为一个抽象层, 在这个层中, 依赖协议的身份验证详细信息对用户隐藏, 要求用户只提供一个 x509 证书。我们建立了 dynafed 的实例, 并将其集成到 atlas 数据分发管理系统中。我们报告了安装和集成过程中面临的挑战。我们测试了 panda

生产系统提交的 atlas 分析作业, 并报告了我们在其运行方面的首次经验。少

2018 年 10 月 26 日提交;最初宣布 2018 年 10 月。

### 3. 理解双面论证在在线消费者评论中的作用: 基于语言的视角

作者: [bernhard lutz](#), [nicolas pröllochs](#), [dirk neumann](#)

**摘要:** 本文考察了双面论证对网上消费者评价感知的有用性的影响。与以前的作品不同的是, 我们的分析从基于语言的角度揭示了评论的接受情况。为此, 我们提出了一种基于分布式文本表示和多实例学习的有趣的文本分析方法, 以实现复习文本中的双面论证。随后使用大量亚马逊评论进行的实证分析表明, 评论中的双面论证显著增加了它们的帮助。我们发现, 积极评价的这种效果比消极评价的效果更强, 而更程度的情感语言会削弱这种效果。我们的发现对零售商平台有直接的影响, 零售商平台可以利用我们的结果来优化他们的客户反馈系统, 并提出更有用的产品评论。少

2018 年 10 月 25 日提交;最初宣布 2018 年 10 月。

### 4. 一种任务包计划程序, 容忍云中的时间故障

作者: [luan teylo](#), [lúcia maria de a.drummond](#), [luciana arantes](#), [pierre sens](#)

**摘要:** ...。机器 (vm) 确保在可用性和波动性方面提供不同的保证, 通过多个定价模型提供相同的资源。例如, 在 **amazon ec2** 云中, 用户为按需虚拟机支付每小时的费用, 而现场虚拟机则是未使用的实例, 价格更低。尽管有货币优势, 但现场虚拟机可以被终止、停止.....。更多

**2018 年 10 月 24 日提交;最初宣布** 2018 年 10 月。

## 5. 一种用于实时多变量优化的高效强盗算法

**作者:** [daniel nhill](#), [houssam nassif](#), [yi liu](#), [anand iyer](#) , [s v n Vishwanathan](#)

**摘要:** 通常使用优化来确定网页的内容, 例如最大限度地提高登录页面上的转换或搜索引擎结果页上的点击率。通常, 这些页面的布局可以分离为几个单独的决策。例如, 着陆页的组成可能涉及决定要显示的图像、要使用的措辞、要显示的颜色背景等。这种优化是一个指数级大决策空间上的组合问题。随机实验不能很好地扩展到此设置, 因此, 实际上, 一个实验通常仅限于一次优化网页的单个方面。这意味着在实验速度和利用布局决策之间可能的交互方面错失了机会。在这里, 我们关注交互式网页的多变量优化。我们制定了一种方法, 在这种方法中, 页面的不同组件之间可能的交互被显式建模。应用土匪方法有效地探索布局空间, 利用爬山实时选择最佳内容。我们的算法还扩展到布局选择的上下文化和个性化。仿真结果表明, 我们的方法适用于内容之间具

有很强交互作用的大型决策空间。我们进一步应用我们的算法来优化促进采用 **amazon** 服务的消息。在仅仅一周的在线优化之后, 我们看到与中位布局相比, 转换增加了 21%。我们的技术目前正在部署, 以优化亚马逊网站多个地点的内容。少

**2018 年 10 月 22 日提交;最初宣布** 2018 年 10 月。

## 6. 在伪装面前捕捉松散的同步行为

**作者:**[潘基文](#), [孙娇](#), [刘欣](#)

**摘要:** 欺诈对社交网络和其他在线应用程序的业务产生了严重的不利影响。用户可以通过在微博上购买 "僵尸追随者" 成为假名人。商家可以通过**亚马逊**上的虚假评论来提升自己的声誉。这种现象在 facebook、yelp 和 tripadvisor 等方面也明显存在。在所有这些情况下, 欺诈者都试图通过伪造他们控制的假账户与目标客户之间的互动来操纵平台的排名机制。少

**2018 年 10 月 20 日提交;最初宣布** 2018 年 10 月。

## 7. cure-or: 挑战不真实和真实环境的对象识别

**作者:**[dogancan temel](#), [jinsol lee](#) , [ghassan alregib](#)

**摘要:** 本文介绍了一种大规模、受控和多平台的对象识别数据集, 该数据集被称为具有挑战性的目标识别的虚幻和真实环境

(cure-or)。在此数据集中, 有 100 个对象的 1, 000, 000 张图像, 这些对象的大小、颜色和纹理各不相同, 它们位于五个不同的方向上, 并使用 5 个设备 (包括网络摄像头、dslr 和三个智能手机摄像机) 在现实世界中 (真实) 和工作室 (不真实) 的环境。可控的具有挑战性的条件包括曝光不足、曝光过度、模糊、对比度、脏镜头、图像噪声、调整大小和丢失颜色信息。我们利用 cure-or 数据集测试 apis-amazon Rekognition 和 microsoft azure 计算机可见性的识别, 并显示它们的性能在具有挑战性的条件下显著降低。此外, 我们还研究了目标识别与图像质量之间的关系, 并表明客观质量算法可以在一定的光度挑战条件下估计识别性能。数据集可在 <https://ghassanalregib.com/cure-or/> 公开使用。

少

**2018 年 10 月 18 日提交;最初宣布** 2018 年 10 月。

## 8. 学习使用文本描述在全局编辑图像

**作者:**王海, [jason d. williams](#), [singbing kang](#)

**摘要:** 我们展示了如何使用文本指令在全局编辑图像: 给定源图像和文本指令进行编辑, 生成在此指令下转换的新图像。为了解决这个新问题, 我们开发了三种不同的基于 mn 和生成对抗性抗网络 (gan) 的可训练模型。模型 (存储桶、筛选器库和端到端) 在编码专家知识的程度上有所不同, 最一般的版本是纯端到端版

本。为了训练这些系统, 我们使用 **amazon** 机械土耳其人收集从多个数据集中采样的大约 2000 个图像对的文本描述。在我们的数据集上评估的实验结果验证了我们的方法。此外, 考虑到滤波器库模型是通用性和性能之间的一个很好的折衷, 我们进一步研究了用图形  $m_n$  替换  $m_n$ , 并证明了图形  $m_n$  提高了性能。据我们所知, 这是第一个纯粹基于自由形式文本指令的全球图像编辑计算摄影作品。少

**2018 年 10 月 12 日提交;最初宣布** 2018 年 10 月。

## 9. ai 辩论中的 30 年周期

作者:[jean-marie chauvet](#)

**摘要:** 在过去的几年里, 人工智能的兴起和这一领域学术突破的成功是不可避免的。向 ai 初创企业投入了大量资金。许多现有的科技公司——包括谷歌、**亚马逊**、脸谱和微软等巨头——已经开设了新的研究实验室。这些日常工作和娱乐工具的迅速变化促使人们对基础技术本身的兴趣不断上升;记者不知疲倦地写 ai, 公司——无论科技性质或非——品牌自己与人工智能, 机器学习或深度学习, 只要他们有机会。面对媒体的这一报道, 几位分析师开始对过度解读 ai 的火爆成功和公众对这一话题的报道有时不佳表示担忧。本文简要回顾了人工智能和机器学习的记录, 发现了早期戏剧性

的成功模式, 其次是哲学批判和意想不到的困难, 如果不是彻头彻尾的停滞, 30 年后几乎回到了时钟自 1958 年以来的周期。少

2018 年 10 月 8 日提交;最初宣布 2018 年 10 月。

## 10. 通过无模型控制提高云计算中的资源弹性

作者: [maria bekcheva](#), [michel Fliess](#), [cédric join](#), [alireza moradi](#), [hugues mounier](#)

摘要: ..... 商用的 "自动缩放" 算法是我们易于实现的方法, 即使在工作负载急剧波动的情况下, 其性能也更好。亚马逊网络服务 (aws) 上的实验证实了这一点。更多

2018 年 10 月 8 日提交;最初宣布 2018 年 10 月。

## 11. 大型网络的有效人群探索: 因果关系的案例

作者: [daniel berenberg](#), [james p. bagrow](#)

摘要: 准确、高效地众包复杂、开放式的任务可能会很困难, 因为人群参与者倾向于支持简短、重复的 "微任务"。我们研究大型网络的众包, 在这些网络中, 人群通过微任务提供网络拓扑。人群可以探索多种类型的社会和信息网络, 但我们关注的是因果属性网络, 这是一个重要的网络, 象征因果关系。我们在 amazon 机械土耳其人 (amt) 上进行实验, 测试工人如何提出和验证个人因果关系, 并引入独立人群工人探索大型网络的方法。该方法的核



心——迭代路径细化, 是通过微观任务进行有效探索的理论原则机制。我们利用合成网络对该方法进行了评价, 并将其应用于 amt 中提取了大规模的因果归因网络, 然后研究了该网络的结构以及构建该网络的工作人员的活动模式和效率。工人的相互作用揭示了因果感知的重要特征, 他们生成的网络数据可以提高我们对因果关系和因果推断的理解。少

2018 年 10 月 7 日提交;最初宣布 2018 年 10 月。

## 12. 快速对象学习和双臂协调, 用于杂乱的发光、拾取和包装

作者 :max schwarz, christian lenz , germán martín garcía, seongyong koo, arul selvam periyasamy, michael schreiber, sven behnke

**摘要:** 从杂乱的垃圾桶里挑选机器人是一项艰巨的任务,亚马逊机器人公司对此提出了挑战。2017 年亚马逊机器人挑战赛 (arc) 要求将物品堆放到存储系统中, 挑选特定的物品, 并将其包装在盒子中。在本文中, 我们描述了 nimbro 选取团队的进入。我们的深度对象感知管道可以快速、高效地适应使用自定义转盘捕获系统和转移学习的新项目。它生产高质量的项目段, 在它被掌握姿势。规划组件协调两个机器人手臂之间的操作操作, 最大限度地减少执行时间。该系统已在 arc 成功演示, 我们的团队在领料任务和最终的问答任务中都获得了第二名。我们还评估各个组件。

少

2018 年 10 月 6 日提交;最初宣布 2018 年 10 月。

### 13. lit: 模型压缩的中间表示培训

作者:[animesh koratana](#), [daniel kang](#), [peter bailis](#), [matei zaharia](#)

**摘要:** 知识蒸馏 (kd) 是一种常用的降低深度网络推理计算开销的方法, 在这种方法中, 教师模型的输出被用来训练一个更小、更快的学生模型。提示训练 (即 fitnet) 通过将学生模型的中间表示性回归到教师模型的中间表示形式来扩展 kd。在本工作中, 我们引入了 block-wise 的中间表示培训 (lit), 这是一种新的模型压缩技术, 它扩展了中间表示在深度网络压缩中的使用, 优于 kd 和提示训练。lit 有两个关键的想法: 1) lit 通过直接比较中间表示来培训与教师宽度相同 (但深度较浅) 的学生; 2) lit 使用教师模型中上一个块中的中间表示作为对当前学生在培训过程中的障碍, 避免了学生网络中不稳定的中间表示。我们表明, lit 提供了大量的网络深度, 而不会损失准确性—例如, lit 可以压缩一个 resnx-110 到一个 ResNeXt-20 (5.5x) 在 cifar10 和 vdcnn-29 到 vdcnn-9 (3.2 x) 在**亚马逊**上的评论, 而不会损失准确性, 性能优于 kd, 并提示网络大小的培训, 以获得给定的准确性。我们还表明, 将 lit 应用于相同的学生/教师架构提高了教师模型高于教师模型的准确性, 优于最近提出的 resnet、resnext 和 vdnnt 上的重生网络程序。最后, 我们证明了 lit 可以有效地

压缩 gan 生成器, kd 框架不支持它, 因为 gan 输出像素而不是概率。少

2018 年 10 月 1 日提交;最初宣布 2018 年 10 月。

#### 14. 多样化的音乐建议

作者 :[houssam nassif](#), [kemal al al cansizlar](#), [mitchell goodman](#), [svn Vishwanathan](#)

**摘要:**我们比较了亚模块化和雅卡德的方法, 使**亚马逊**音乐的建议多样化。子模块化显著提高了推荐质量和用户参与度。与 jacard 方法不同, 我们的子模块化方法将项目相关性评分纳入其优化功能, 并生成一个相关且一致多样化的集合。

2018 年 10 月 2 日提交;最初宣布 2018 年 10 月。

#### 15. 用于视觉发现的自适应、个性化的多样性

作者 :[choon hui teo](#), [houssam nassif](#), [daniel hill](#), [sriram srinivasan](#), [mitchell goodman](#), [vijai mohan](#), [svn Vishwanathan](#)

**摘要:** 当用户有明确的意图时, 搜索查询是合适的, 但当意图难以表达或用户只是希望受到启发时, 它们的性能很差。可视化浏览系统允许电子商务平台解决这些方案, 同时为用户提供引人入胜的购物体验。在这里, 我们探索流中自适应个性化和项目多样化的方向的扩展, 这是亚马逊视觉浏览和发现的一种新形式。我

们的系统为用户提供了一组不同的有趣项目，同时适应了用户交互。我们的解决方案由三个组件组成 (1) 贝叶斯回归模型，用于在利用不确定性的同时对项目的相关性进行评分; (2) 根据类别对得分最高的项目进行重新排序的子模块多样化框架; (3) 个性化从用户的行为中了解到的类别首选项。在实时流量上进行测试时，我们的算法显示在点击率和会话持续时间上有很强的提升。

少

2018 年 10 月 2 日提交;最初宣布 2018 年 10 月。

## 16. 继电器: 一种适用于机器学习框架的新 ir

作者 : [jared roesch](#), [steven lyubomirsky](#), [logan weber](#), [josh pollock](#), [marisa kirisame](#), [tiqichen](#) , [zachary tatlock](#)

**摘要:** 机器学习为工业中的各种服务提供动力，包括搜索、翻译、推荐系统和安全。这些模型的规模和重要性要求它们在一系列异构硬件设备中高效、富有表现力和便携。这些制约因素往往是不一致的;为了更好地适应它们，我们提出了一种新的高级中间表示 (ir)，称为中继。中继被设计为一种纯功能的静态类型的语言，其目标是平衡高效编译、表现力和可移植性。我们讨论了中继的目标，并强调了它的重要设计限制。我们的原型是开源 `nnvm` 编译器框架的一部分，它为 **amazon** 的深度学习框架 `mxnet` 提供了动力。

少

2018 年 9 月 25 日提交;最初宣布 2018 年 10 月。

## 17. rgb-d 在杂波中自主操作的 rgb-d 目标检测与语义分割

作者: [max schwarz](#), [anton milan](#), [arul selvam periyasamy](#), [sven behnke](#)

**摘要:** 杂波中的自主机器人操作具有挑战性。必须在复杂的场景中感知到各种各样的对象, 这些对象被部分遮挡并嵌入到许多干扰体中, 通常位于受限空间中。为了应对这些挑战, 我们开发了一种将对象检测和语义分割结合起来的深度学习方法。操作场景是用 rgb-d 摄像机拍摄的, 为此我们开发了一种深度融合方法。采用预先培训的功能, 可以从小的带注释的机器人数据集中学习。我们在两个具有挑战性的数据集上评估我们的方法: 一个是为 2016 年 **亚马逊** 采摘挑战赛捕获的, 我们的团队 **nimb 罗** 在 "stowing" 中名列第二, 在拾取任务中排名第三, 另一个在灾难场景中捕获。实验表明, 目标检测和语义分割是相辅相成的, 可以结合起来, 得到可靠的目标感知。少

2018 年 10 月 1 日提交;最初宣布 2018 年 10 月。

## 18. gdpr 后的隐私政策格局

作者: [thomas linden](#), [hamza harkous](#), [kassem fawaz](#)

**摘要:** 每一个新的隐私规定都会带来这样的问题: 它是否会改善用户的隐私, 或者是否会对理解和行使他们的权利造成更多的障碍。欧盟一般数据保护条例 (gdpr) 是有史以来要求最高、最全面的隐私法规之一。因此, 在它生效几个月后, 自然会研究它对网上隐私政策格局的影响。在这项工作中, 我们对 gdpr 前后的隐私政策进行了第一次纵向、深入和规模的评估。我们衡量这些政策的完整消费周期, 从第一次用户印象到合规性评估。我们创建了一个由 3 086 英语隐私政策组成的多样化语位, 并为其获取 gpdr 前版本和 gdpr 后版本。通过 **amazon** mturk 上 530 参与者的用户研究, 我们发现, 除了欧洲顶级网站外, 隐私政策的可视化呈现在有限的数据敏感类别中略有改进。我们还发现, 隐私政策的可读性受到 gdpr 的影响, 因为尽管为减少对被动判决的依赖做出了努力, 但判决和词语增加了近 3 0%。我们在自动化自然语言处理技术的基础上, 进一步开发了一种新的工作流程, 用于自动评估隐私政策中的要求。我们发现了 gdpr 引发积极变化的证据, 模糊程度模糊, 平均超过 8 个指标, 超过 20.5% 的政策有所改善。最后, 我们展示了隐私策略涵盖了更多的数据实践, 特别是围绕数据保留、用户访问权限和特定受众, 平均 15.2% 的策略在 8 个合规性指标中得到了改进。然而, 我们的分析揭示了目前的现状与 gdpr 的最终目标之间存在着很大的差距。少

**2018 年 9 月 22 日提交;最初宣布 2018 年 9 月。**

## 19. 利用离岸论证挖掘预测亚马逊评论的有用性

作者:marco passon, marco lippi, giuseppe serra, carlo tasso

**摘要:** 互联网用户以前所未有的速度生成内容。因此, 建立能够在这—信息海洋中区分有用内容的智能系统正成为迫切需要。在本文中, 我们的目标是预测 **amazon** 评论的有用性, 为此, 我们利用了来自现成论证挖掘系统的功能。我们认为, 事实上, 评审的有用性与其论证内容严格相关, 而使用已经受过训练的系统则避免了重新标记新数据集的成本高昂的需要。在一个大的公开语料库上获得的结果支持了这一假设。少

2018 年 9 月 21 日提交;最初宣布 2018 年 9 月。

## 20. 分析谷歌上的社交图书阅读行为及其对亚马逊最佳卖家的预测

作者 :suan kalyan maity, abhishek panigrahi, animesh mukherjee

**摘要:** 一本书的成功程度取决于各种参数——外在和内在的。本文研究了图书阅读特性如何影响一本书的普及程度。为此, 我们对 goodreads 实体进行了跨平台研究, 并试图在 goodreads 的各个实体和热门图书 ("**亚马逊**畅销书") 之间建立联系。我们分析了 goodreads 平台上的集体阅读行为, 并量化了 goodreads 实体的各种特征, 以识别这些**亚马逊**畅销书 (abs) 与其他非畅销图书之间的差异。然后, 我们开发了一个预测模型, 利用特征特征

来预测一本书自出版一个月（15天）后是否会成为畅销书。在平衡的集合中，我们能够实现非常高的平均精度为 88.72（含 10.66%）的预测，其中其他竞争类包含从 goodreads 数据集中随机选择的书籍。我们的方法主要基于从用户帖子和体裁相关特征属性中获得的特征，比传统的流行因素（评级、评论）基于基线方法提高了 16.4%。我们还评估我们的模型与两个更有竞争力的书籍集 a) 都是高评价，并已收到了大量的评论（但不是畅销书）（人权保护）和 b) 谷歌选择奖提名书籍是非畅销书（gcan）。我们能够以非常高的平均精度 87.1% 和高 roc 的 abs vs. gcan 取得相当好的结果。对于 abs 与 hrhr 相比，我们的模型的平均精度为 86.22%。少

2018 年 9 月 19 日提交;最初宣布 2018 年 9 月。

## 21. 基于多语言神经机器翻译的零射击跨语言分类

作者 :[akiko eriguchi](#), [melvin johnson](#), [orhan firat](#) , [hidto kazawa](#), [wolfgang macherey](#)

摘要: 将表示从大型监督任务转移到下游任务在人工智能领域(如计算机视觉和自然语言处理(nlp))中显示出很有希望的结果。与此同时,机器翻译(mt)的最新进展使人们能够培训多语言神经 mt (nmt) 系统,这些系统可以在多种语言之间进行翻译,并且还能够执行零镜头翻译。但是,很少注意利用多语言 nmt 系统所学到的表示,以便在其他 nlp 任务中实现零拍摄多语言性。在



本文中，我们演示了一个简单的框架，多语言编码器分类器，通过重用编码器从多语言 nmt 系统，并将其与任务特定的分类器组件拼接，进行跨语言的跨语言迁移学习。我们提出的模型在三个基准任务（amazon review、sst 和 snli）的英语设置方面实现了显著改进。此外，我们的系统可以使用在培训期间看不到分类数据的新语言进行分类，这表明零镜头分类是可能的，而且竞争非常激烈。为了了解促成这一发现的潜在因素，我们对共享词汇的影响、nmt 的训练数据类型、分类器的复杂性、编码器表示能力和模型泛化进行了一系列分析。零拍摄性能。我们的研究结果提供了有力的证据，证明从多语言 nmt 系统中学到的陈述广泛适用于各种语言和任务。少

2018 年 9 月 12 日提交;最初宣布 2018 年 9 月。

## 22. 动态定价竞争中的维度曲线处理：利用频繁的定价来补偿市场预期不完善的问题

作者:[rainer schlosser](#), [martin boissier](#)

**摘要:** 大多数销售应用程序的特点是竞争和有限的需求信息。对于成功的定价策略，频繁的价格调整以及对市场动态的预期至关重要。这两种影响都具有挑战性，因为竞争激烈的市场是复杂的，优化的定价调整的计算可能非常耗时。分析了易腐商品销售寡头垄断竞争下的随机动态定价模型。为了避免维度的诅咒，我们提出

了一种启发式方法来有效地计算价格调整。为了证明我们的战略的适用性,即使竞争对手的数量很大,他们的战略是未知的,我们考虑不同的竞争环境,其中竞争对手频繁和战略性地调整他们的价格。对于所有设置,我们验证我们的启发式策略是否产生了有希望的结果。我们将我们的启发式与上限的性能进行比较,这些启发式是通过利用完美的价格预期的最优策略获得的。我们发现,价格调整频率对预期利润的影响比预期的价格更大。最后,我们的方法已应用于**亚马逊**的旧书籍的销售。我们用卖家的历史市场数据来校准我们的模型。销售结果显示,我们的数据驱动战略的业绩优于有经验的卖家基于规则的战略,利润增长超过 20%。

少

2018 年 9 月 7 日提交;最初宣布 2018 年 9 月。

## 23. 利用深度学习从亚马逊雨林卫星图像中识别土地模式

作者:[somnath rakshit](#), [soumyadeep Somnath](#), [dhiman mondal](#)

**摘要:** **亚马逊**雨林一直遭受广泛的破坏,无论是通过自然和人工手段。每一分钟,估计世界失去森林覆盖 48 个足球场的大小。**亚马逊**雨林的森林砍伐导致生物多样性急剧减少、栖息地丧失、气候变化和其他生物损失。在这方面,必须跟踪这些森林的性质如何随着时间的推移而变化。使用深度学习进行图像分类可以通过删除对每个图像进行分类的手动任务来帮助加快这一过程。在这

里，它展示了如何卷积神经网络可以用来跟踪**亚马逊**雨林的土地模式的变化。在本工作中，测试精度为 967.1%。这可以帮助各国政府和其他机构更有效和准确地跟踪土地格局的变化。少

2018 年 9 月 2 日提交;最初宣布 2018 年 9 月。

## 24. 支付一，免费获得数百：通过共享查询执行降低云成本

作者 :[renato marroquín](#), [ingo müller](#), [darko makreshanski](#), [gustavo alonso](#)

**摘要:** 基于云的数据分析是当今的普遍做法，因为系统管理开销较低，以及现收现付定价模型。然而，定价模型并不总是适合查询处理，因为大量使用会导致高昂的成本。例如，在查询即服务系统中，用户按处理的字节付费，频繁访问相同数据的查询集合可能会变得昂贵。在使用声明性接口（如 sql）时，用户优化查询执行的选项有限，使问题更加复杂。在本文中，我们展示了如何在不修改现有系统和云提供商的参与的情况下，显著降低查询即服务系统的开销，从而降低成本。我们的方法基于查询重写，以便将多个并发查询合并到一个查询中。我们的实验表明，共享执行所完成的工作总量小于一次查询方法。由于查询是按处理的每个字节计费的，因此执行一组查询的成本通常与执行其中的一个查询的成本相同。例如，我们演示了在 amazon athena 和 google bigquery 中，与在**实现更高吞吐量**的同时使用实时查询方法相比，tpc-h 基准的共享执行成本如何高达 100x 和 16x。少

2018 年 9 月 1 日提交;最初宣布 2018 年 9 月。

## 25. 从只使用语音的界面中获得的早期教训

作者:[joshua wissbroecker](#), [f maxwell harper](#)

摘要: ..... 除了语音, 没有输入模式的系统。在这张海报中, 我们描述了电影电视, 重点介绍了在现成的 **amazon echo** 周围构建原型系统的经验教训。更多

2018 年 8 月 29 日提交;最初宣布 2018 年 8 月。

## 26. 网络无人机防御群对恶意无人机入侵的防御

作者:[matthias r. brust](#) , [grégoire d 赛尼](#), [pascal bouvry](#), [dren gashi](#), [himadripathak](#), [mike p. gonçaves](#)

摘要: 如今,**亚马逊**、阿里巴巴甚至比萨饼连锁店等公司都在推动使用无人机 (也称为无人驾驶飞行器), 以提供服务, 如提供套餐和食品。由于各国政府打算利用无人机所能提供的这些巨大的经济利益, 城市规划人员正在着手将所谓的无人驾驶飞行器飞行区和无人机高速公路纳入其智能城市设计。然而, 需要监视无人机的高速移动性和行为动态, 以检测并随后处理具有恶意的入侵者、流氓无人机和无人机。本文提出了一种无人机防御系统, 用于拦截和护送飞行区外的恶意无人机。提出的无人机防御系统由防御无人机群组成, 该群能够在入侵者检测的情况下自行组织其防御

编队, 并将恶意无人机作为网络化群追逐。模块化设计原则已用于我们完全本地化的方法。我们开发了一个创新的自动平衡聚类过程, 以实现拦截和捕获的形成。事实证明, 由此产生的网络防御无人机群对通信损失具有弹性。最后, 实现了无人机仿真样机。通过大量的模拟, 我们展示了我们的方法的可行性和性能。少

**2018 年 9 月 2 日提交;**v1 于 2018 年 8 月 21 日提交;**最初宣布** 2018 年 8 月。

## 27.lrmm: 学习推荐与缺失的方式

**作者:**王成,马蒂亚斯·涅特珀特,李辉

**摘要:** 多模式学习在基于内容的推荐中表现出了很有希望的性能, 因为文本和图像等多种模式的辅助用户和项目信息。然而, 不完全和缺失模式的问题很少被探讨, 大多数现有方法未能学习模式缺失或损坏的推荐模式。在本文中, 我们提出了一个新的框架, 不仅可以缓解缺少模式的问题, 而且可以更普遍地缓解推荐系统的冷启动问题。我们建议模态降 (m-下降) 和多模态顺序自动编码器 (m-drop), 以学习多模态表示, 以补充和估算缺失的模式。在实际**亚马逊**数据上进行的大量实验表明, lrmm 在评级预测任务方面实现了最先进的性能。更重要的是, lrmm 在缓解数据稀疏性和冷启动问题方面对以前的方法更加强大。少

**2018 年 8 月 30 日提交;v1 于 2018 年 8 月 21 日提交;最初宣布 2018 年 8 月。**

## **28. 一种可扩展的双足机器人：平面模型中的可变步态合成与验证**

**作者:**[umer huzafa](#), [catherine maguire](#), [amy laviers](#)

**摘要:** 人类是高效的,但在他们的运动中富有表现力。人类行走行为可以用来在不坠落的情况下行走,并通过不同的步态风格与他人沟通内部状态。这为创建类似的富有表现力的双足机器人提供了灵感。为此,提出了一个类似于复合驱动的平面双足模型中的风格步态生成框架。步态设计是利用基于模型的具有可变约束的轨迹优化完成的。对于有限范围的优化参数,可以为该模型生成大量 360 步态。特别是步长和成本函数的变化,产生了明显的循环行走步态。从这些产生的步态,6 个步态被识别和标记,使用体现运动分析,与人类活动相关的文体动词,如 "lope" 和 "saunter"。通过在 **amazon** 机械土耳其语中进行用户研究,这些标签已经得到了验证,从而证明使用此框架可以生成视觉上可区分的、有意义的步态。这为创建具有可变社会能力的运动特征的双足人形奠定了基础。少

**2018 年 9 月 13 日提交;v1 于 2018 年 8 月 16 日提交;最初宣布 2018 年 8 月。**

## **29. 向非专家解释 web 表查询**

**作者:**jonathan berant, daniel gerch, amirgloberson, tova milo, tomer wolfson

**摘要:** 为查询表设计可靠的自然语言 (nl) 接口一直是数据管理和自然语言处理 (nlp) 社区研究人员的长期目标。这样的接口接收作为输入 nl 问题, 将其转换为正式查询, 执行查询并返回结果。翻译过程中的错误并不少见, 用户通常难以理解其查询是否已正确映射。我们通过向非专家用户解释所获得的正式查询来解决此问题。提出了两种查询解释方法: 第一种方法将查询转换为 nl, 而第二种方法提供基于查询单元的源的图形表示形式 (在给定表中的执行中)。我们的解决方案通过 web 表增强了最先进的 nl 接口, 在培训和部署阶段都增强了该接口。实验, 包括在 amazon 机械土耳其人身上进行的用户研究, 展示了我们提高 nl 接口正确性和可靠性的解决方案。少

**2018 年 8 月 14 日提交;最初宣布** 2018 年 8 月。

### 30. 用于推断配对溢出效应的链接原因变分自动编码器

**作者:**vineeth rakesh, ruocheng guo, raha moraffah , nitin Agarwal, huan liu

**摘要:** 来自观测数据的建模溢出效应是经济学、商业和其他研究领域的一个重要问题。% 它帮助我们推断两个看似无关的事件集之间的因果关系。例如, 如果美国的消费支出下降, 就会对依赖美国作为其最大出口市场的经济体产生溢出效应。在本文中, 我们的

目的是推断导致对实体（或单位）之间的溢出效应的原因，我们称之为 \text{成对溢出}。为了实现这一目标，我们利用近年来在变分推理和深度学习技术方面的发展，提出了一种称为链接因果变化自编码器（lcva）的生成模型。与变分自动编码器（vae）类似，lcva 采用了编码器神经网络来学习潜在属性和解码器网络来重建输入。但是，与 vae 不同的是，lcva 将无级别标题 {潜在属性视为假定同时影响处理和单位结果} 的混淆属性。具体来说，给定一对单位美国和美国<sup>-</sup>，他们的个人处理和结果，编码器网络的 lcva 采样的混淆通过对观察到的协变美国，治疗这两种方法美国和美国<sup>-</sup>和结果美国。一旦推断，潜在的属性（或混淆）美国捕捉溢出效应美国<sup>-</sup>上美国。利用来自作业培训数据集（Ilonde (1986)）的用户网络和亚马逊电子商务领域的共同购买数据集，我们表明，lcva 在捕获溢出效应方面比现有方法更加强大。少

**2018 年 10 月 3 日提交**;v1 于 2018 年 8 月 9 日提交;**最初宣布** 2018 年 8 月。

### 31. 可控制的图像到视频的翻译——以面部表情生成为例

**作者:**范丽杰,黄文兵,庄甘,黄俊洲,龚博清

**摘要:** 最近在深度学习方面的进步使得利用神经网络生成照片逼真图像成为可能，甚至可以从输入视频剪辑中推断视频帧。本文研究了图像到视频的翻译，特别是面部表情的视频，既是为了进



一步探索这一探索，也是为了促进我们自己对现实应用的兴趣。与图像到图像的平移相比，这个问题通过另一个时间维数来挑战深度神经网络。此外，它的单个输入图像无法使用大多数依赖于重复模型的现有视频生成方法。我们提出了一种用户可控的方法，以便从单个人脸图像生成不同长度的视频剪辑。表达式的长度和类型由用户控制。为此，我们设计了一种新的神经网络体系结构，该体系结构可以将用户输入集成到其跳过连接中，并对神经网络的对抗训练方法提出了一些改进建议。实验和用户研究验证了我们的方法的有效性。特别是，我们要强调的是，即使是野外的面部图像（从网上下载和作者自己的照片），我们的模型可以生成高品质的面部表情视频，其中约 50\% 被[亚马逊](#)标记为真实土耳其机械工人。少

**2018 年 8 月 8 日提交;最初宣布** 2018 年 8 月。

### 32. 大规模语言建模: 在 4 小时内收敛到 40gb 的文本

**作者:**[劳尔·普里](#),[罗伯特·柯比](#),[尼古拉·雅科文科](#),[布莱恩·卡坦扎罗](#)

**摘要:** 最近的工作表明了如何在大型图像数据集上快速训练卷积神经网络 (cnn)，然后将从这些模型中获得的知识转化为各种任务。在 [radford 2017] 之后，在本工作中，我们展示了用于自然语言任务的重复神经网络 (mn) 的类似可扩展性和传输。通过使用分布在 128 nvidia tesla v100 gpu 的混合精度算法和 32k 批处

理大小, 我们能够在 4096 维乘法 lstm (mlstm) 上训练一个字符级 4096 维乘法 lstm (mlstm), 用于 40 gb 的 3 个世纪的无监督文本重建亚马逊在四个小时内查看数据集。与以前需要一个月的时间在同一数据集中为一个时代训练相同的大小和配置的工作相比, 此运行时是有利的。融合大批量 mn 模型可能具有挑战性。最近的工作建议将学习速率扩展为批处理大小的函数, 但我们发现, 简单地将学习速率扩展为批处理大小的函数, 就会导致此问题的收敛性明显恶化或立即出现分歧。我们提供了一个学习速率计划, 允许我们的模型收敛到 32k 批处理大小。由于我们的模型在数小时内就会在 amazon 评审数据集上进行扩展, 而我们对 128 tesla v100 gpu 的计算要求虽然在商业上是大量的, 但这项工作为大多数商业领域提供了大规模的无监督 nlp 培训应用和深入的学习研究人员。模型可以在任何公共或私有文本数据集上进行一夜之间的训练。少

2018 年 8 月 10 日提交;v1 于 2018 年 8 月 3 日提交;最初宣布 2018 年 8 月。

### 33. 深入了解是什么让科学可视化令人难忘

作者:李瑞,陈健

摘要: 我们报告了一项初步研究的结果, 该研究探讨了空间科学可视化的记忆能力, 其目的是了解有助于记忆能力的视觉特征。

评估指标包括三个客观度量 (熵、特征拥塞、边缘数)、四个主观评级 (杂乱、不同颜色的数量、熟悉程度和现实主义) 和两个情绪评级 (有趣和快乐).我们从 2008 年至 2017 年出版 *ieee svvis* 论文中的原始 2231 张图像中收集了 1142 个科学可视化 (svis) 图像, 并从[亚马逊机械土耳其 \(mturk\)](#) 上收集的数据中计算了 228 个 svvis 图像的记忆分数。结果表明, scivis 图像的记忆能力主要与杂波和不同颜色的数量有关。我们进一步研究科学可视化和信息图表之间的差异, 作为通过数据属性理解记忆差异的一种手段。

少

**2018 年 8 月 26 日提交;v1** 于 2018 年 8 月 1 日提交;**最初宣布** 2018 年 8 月。

### 34. 电子商务赞助搜索广告中拍卖机制的实践约束优化

作者:[白刚](#),[谢志辉](#),[王亮](#)

**摘要:** 在[亚马逊](#)、[淘宝](#)和[天猫](#)等电子商务平台上进行赞助搜索, 为卖家提供了一个有效的方式, 以达到最相关的目的。本文研究了阿里巴巴移动电子商务平台上赞助搜索中的拍卖机制优化问题。除了创造收入, 我们应该保持一个高效的市场, 有大量的优质用户, 保证广告商的合理投资回报 (roi), 同时, 促进用户愉快的购物体验。这些要求本质上是一个受约束的优化问题。直接优化拍卖参数会产生一个不连续的、非凸的问题, 从而否定有效的解决

方案。我们的主要贡献之一是对原始问题进行实际的凸优化配方。我们设计了一种新的具有代表性实例的离散检验机制的再参数化方法。为了构造优化问题, 我们建立了一个拍卖仿真系统, 通过重播真实在线请求记录的拍卖, 估计所选参数的结果业务指标。总结了真实搜索流量的实验, 分析了拍卖仿真的保真度、各种约束目标下的效果以及正则化的影响。实验结果表明, 通过适当的熵正则化, 我们能够在一定范围内最大限度地提高收入, 同时约束其他业务指标。少

**2018 年 7 月 31 日提交;最初宣布** 2018 年 7 月。

### 35. 在 lte 中实现多播流的最优分组与资源分配

**作者:**[sadaf ul zuhra](#), [prasanna chaporkar](#), [abhay karandikar](#)

**摘要:** ..... 预计到 2020 年, 流量将占总数据流量的 82%。随着 youtube、netflix、amazon prime video 等视频流应用程序的日益普及, 流行的视频内容通常需要同时传递给大量用户。多播传输可用于满足此类应用。更多

**2018 年 7 月 30 日提交;最初宣布** 2018 年 7 月。

### 36. 生产无服务器功能编排系统的比较

**作者 :**[pedro garcía lópez](#), [marc sánchez-artezas](#), [gerard parís](#), [daniel barshpons](#), [alvaro ruiz ollo](#) 丁·[david arroyo pinto](#)

**摘要:** 自 2014 年 **amazon** **lambda** 出现以来, 所有主要云提供商都采用了 "功能即服务 (faas)" 模型, 因为它具有巨大的潜力, 可用于各种应用。正如预期的那样 (也是所希望的), 在无服务器世界中竞争非常激烈, 包括对无服务器函数的编排的运行时支持等方面。在这方面, 目前的三个主要参与者是**亚马逊**步进函数 (2016 年 12 月)、**azure** 持久函数 (2017 年 6 月) 和 **ibm** 作曲家 (2017 年 10 月), 这些项目仍然是年轻的实验项目, 还有很长的路要走。在本文中, 我们将在一个通用的评估框架下比较和分析这三个无服务器业务流程服务。我们将研究它们的体系结构、编程和计费模型, 以及它们对并行执行的有效支持等。通过一系列实验, 我们还将评估不同类型工作流的不同基础结构的运行时开销。

少

2018 年 7 月 30 日提交;最初宣布 2018 年 7 月。

### 37.位置、位置、位置: 探索跨地理区域的 **amazon ec2** 现货实例定价-扩展版本

**作者:**[nnamdi ekwe-ekwe](#), [adam barker](#)

**摘要:** 云计算正在成为计算领域中几乎无处不在的一部分。对于当今的许多公司来说, 将其整个基础架构和工作负载迁移到云中可以降低复杂性、部署时间并节省资金。竞价型实例是 **amazon** 云计算基础架构 (**ec2**) 的一个子集, 它对此进行了扩展。它们允许

用户以大幅打折的价格对**亚马逊**数据中心的备用计算能力进行竞价。如果需求不断增加,从而超过了用户的最高出价,则会终止其实例。在本文中,我们对位置如何影响现场实例的总体部署成本进行了首次详细分析之一。我们分析所有可用的 **amazon** web 服务区域的定价数据,为期 60 天,用于各种现场实例类型。我们将发现的数据与整个 **aws** 区域以及该区域内的可用区联系起来。我们得出的结论是,位置在现场实例定价中确实发挥着至关重要的作用,而且定价也因位置的粒度而异—从更粗粒度的 **aws** 区域到区域内更细粒度的可用区。我们将我们发现的价格差异与价格的可靠性联系起来,确认一个人是否可以对报告的价格和随后的出价充满信心,在随后的出价中。最后,我们表明,在竞价型实例上运行工作负载是可能的,既可以降低终止风险,也可以实现每小时支付非常低的金额。少

**2018 年 7 月 27 日提交;最初宣布** 2018 年 7 月。

### 38. 2p-dn:隐毛深部神经网络 (基于同态密码系统)

作者:[朱强](#),[吕锡祥](#)

**摘要:** 机器学习即服务 (mlaas) (如 **microsoft azure**、**amazon aws**) 提供了一个有效的 dnn 模型,可以为仅限于缺乏数据和计算能力的小型企业和个人完成机器学习任务。但是,这里出现了一个问题,即用户隐私被提交给 mlaas 服务

器, 因为用户需要将其敏感数据上传到 mlaas 服务器。为了保护他们的隐私, 用户可以在上传数据之前对其进行加密。这使得运行 dnn 模型变得很困难, 因为它不是为在密文域中运行而设计的。本文利用 paillier 同态密码系统, 提出了一种新的保留强效神经网络模型, 称为 2p-dnn。该模型可以完成密文域中的机器学习任务。通过使用 2p-dnn, mlaas 能够为用户提供一个预置保护机器学习服务。我们基于 lenet-5 构建 2p-dnn 模型, 并使用加密的 mnist 数据集对其进行测试。分类精度在 97% 以上, 接近于与 mnist 数据集一起运行的 lenet-5 的准确率, 并且高于其他现有的优先机器学习模型

2018 年 7 月 23 日提交;最初宣布 2018 年 7 月。

### 39. 相等但不相同: 理解可说服图像与文本之间的隐式关系

作者:张明达, [rebecca hwa](#) , [adriana kovashka](#)

**摘要:** 广告中的图像和文本以复杂的非文字方式相互作用。这两个频道通常是互补的, 每个频道讲述的故事都是不同的部分。当前的方法, 如图像字幕方法, 只检查文字, 冗余关系, 其中图像和文本显示完全相同的内容。为了了解更复杂的关系, 我们首先收集广告解释的数据集, 以确定同一视觉广告中的图像和口号是否形成并行 (传达相同的信息而不说出相同的内容) 或在亚马逊机械土耳其人招聘的工人的帮助下, 建立了不平行的关系。我们开发

了各种功能, 以捕捉图像的创造力和文本的特殊性或模糊性, 以及分析渠道内和渠道之间的语义的方法。在预测图像与文本的并行性关系时, 我们的方法优于标准的图像-文本对齐方法。少

**2018 年 7 月 21 日提交;最初宣布** 2018 年 7 月。

#### 40. 智能空间中基于对话的复杂事件管理

**作者:**[andrésousa lago](#), [hugo sereno ferreira](#)

**摘要:** 智能空间管理可以通过多种方式完成。一方面, 有对话助理, 如谷歌助理或**亚马逊**亚历克莎, 使用户能够舒适地与智能空间与智能空间与他们的声音, 但这些功能有限, 通常仅限于简单命令。另一方面, 有一些可视化的接口, 如 ibm 的 node-red, 可在不同设备之间启用复杂的功能和依赖关系。然而, 这些都是有限的, 因为它们要求用户对智能设备的工作方式有技术知识, 而且系统的接口更复杂, 更难以使用, 因为他们需要一台计算机。该项目提出了一个新的对话助理-贾维斯-结合了当前助理的易用性和视觉平台的操作复杂性。jarvis 的目标是通过提供直观的命令和有用的功能, 使管理智能空间变得更加容易。jarvis 与现有的用户界面 (如 google 助手、slack 或 facebook messenger) 集成, 使其能够很容易地与现有系统集成。贾维斯还提供了创新的功能——因果关系查询——使用户能够问它为什么会发生一些事情。



例如, 用户可以询问 "为什么开灯", 以了解系统是如何工作的。

少

2018 年 7 月 18 日提交;最初宣布 2018 年 7 月。

#### 41. 通过综合审查改进可解释的建议

作者:sixun ouyang, aonghus lawlor, felipe costa, peter dolog

**摘要:** 推荐系统的一项重要任务是为用户提供可解释的解释。这对该系统的信誉很重要。目前可解释的推荐系统倾向于关注已知对用户很重要的某些功能, 并以结构化的形式提供他们的解释。众所周知, 用户生成的审阅和审阅者的反馈对用户的决策具有很强的影响力。另一方面, 最近的文本生成作品已被证明生成的文本质量与人类的书面文本相似, 我们的目的是表明生成的文本可以成功地用于解释建议。在本文中, 我们提出了一个由流行的评论导向生成模型组成的框架, 旨在为建议创建个性化的解释。解释是在字符和单词级别上生成的。我们构建一个数据集, 其中包含来自**亚马逊**图书评审数据集的审阅者的反馈。我们的跨域实验旨在将自然语言处理与推荐系统领域之间的联系起来。除了语言模型评价方法外, 我们还采用了一种新型的基于深度神经网络的面向评审的推荐系统 **deepcon**, 通过根均方误差 (rmse) 来评价生成的评审的推荐性能。我们证明, 综合个性化的评论比人类的书

面评论有更好的推荐性能。据我们所知, 这提供了第一个机器生成的自然选择解释评级预测。少

2018 年 7 月 18 日提交;最初宣布 2018 年 7 月。

## 42. 安全心理模型: 认知地图方法

作者:[tahani albalawi](#), [kambiz ghazinour](#), [austin melton](#)

**摘要:** 安全模型的设计是为了确保根据安全策略以适当的方式访问和使用数据。不幸的是, 人在设计安全模型时的作用被忽略了。人类行为与许多安全漏洞有关, 在许多安全情况下发挥着重要作用。本文通过心理模型的方法, 研究了用户对安全和可用性的安全决策。为了获得和描述用户的安全和可用性心理模型, 采用了人群采购技术和认知地图方法, 我们已经进行了一个实验, 以评估我们的发现使用 **amazon** mturk。少

2018 年 7 月 17 日提交;最初宣布 2018 年 7 月。

## 43. 可解释性建议的分层相关性传播

作者:[homanga bharadhwaj](#)

**摘要:** 在本文中, 我们利用分层关联传播技术, 在基于深度学习的建议模型中了解释问题。在识别特征空间中图像之间的相似性之前, 我们使用深卷化神经网络从输入图像中提取相关特征。

图像之间的关系由模型标识, 并使用层级相关性传播来推断可能显著影响模型选择的图像的像素级详细信息。我们在 **amazon** 产品数据集中评估我们的方法, 并展示我们的方法的有效性。少

**2018 年 7 月 16 日提交;最初宣布** 2018 年 7 月。

#### 44. 街头感觉: 从谷歌街景中学习

**作者:**[suriyan laohaprapanon](#),[金伯利·奥特莱布](#), [gaurav sood](#)

**摘要:** 公共服务和公共基础设施有多好? 他们的素质因收入而异吗? 这些都是至关重要的问题——它们揭示了政府的工作做得有多好, 地方资金差距的后果等。但在这些问题中, 几乎没有什么好的数据。我们通过描述一种可扩展的方法来填补这一空白, 该方法可以在一个关键的公共基础设施上获得数据: 道路。我们利用 google 街景中的数据来评估道路和人行道的质量。我们随机采样主要道路上的位置, 查询这些位置的谷歌街景图像, 并使用**亚马逊的机械土耳其语**对图像进行编码。我们采用这种方法来评估曼谷、雅加达、拉各斯和密歇根州韦恩县的道路质量。雅加达的道路坑洞几乎是其他城市道路的四倍。令人惊讶的是, 曼谷、拉各斯和韦恩的坑洞路段比例大体相当, 在. 06 至 0.07 之间。利用这些数据, 我们还估计了密歇根州韦恩市的路况与当地收入之间的关系。我们发现, 在较富裕的人口普查区, 道路的坑洞较少一些。少

2018 年 7 月 10 日提交;最初宣布 2018 年 7 月。

#### 45. 在您购物时的窃听: 平衡互联零售空间中的个性化和隐私

作者:[vasilios mavroudis](#), [michael veale](#)

**摘要:** 曾经在追踪忠诚卡和 "反向追加" 方面走在前列的实体零售商, 现在落后于在线竞争对手。然而, 我们可能会看到这些表格的转向, 因为许多人越来越多地部署技术, 从简单的传感器到先进的情感检测系统, 甚至使他们能够根据每个客户量身定制价格和购物体验。在这里, 我们将在零售环境中检查这些店内跟踪技术, 并从技术和监管角度对其进行评估。我们首先介绍相关的技术的背景下, 在考虑隐私影响, 目前的补救措施, 个人可能通过技术和法律, 以及这些补救措施的限制。为了说明在这一领域具有挑战性的紧张关系, 我们考虑了对 a) 最近来自**亚马逊**的 "go" 商店概念采取技术和法律方法的可行性, 该概念要求细粒度的多模式跟踪作为商店发挥作用, b) 当前在选择进入或退出日益普及的被动 wi-fi 跟踪方面的挑战。"go" 商店在欧洲的合法性明显不明确, 是单方面的, 以避免生物识别跟踪可能无效, 这给它带来了重大挑战。就 mac 地址而言, 我们看到隐私保密和隐私控制之间难以调和的冲突, 并提出了一个可能有助于平衡两者的技术框架。在寻求在个性化与隐私之间取得平衡时, 存在着重大挑战, 研究人员必须共同努力, 包括跨越首选隐私定义的界限, 以提出既借鉴技术又借鉴法律的解决方案提供有效和相称的保护。

零售商同时必须确保他们的跟踪不仅是合法的, 而且值得有关数据主体的信任。少

2018 年 7 月 14 日提交;最初宣布 2018 年 7 月。

#### 46. 科学令牌: 基于能力的远程科学数据安全访问

作者 :alex withers, brian bockelman, derek weitzel, duncan brown, jeff gaynor, jim basney, todd tannenbaum, zach miller

**摘要:** 计算科学工作流的安全凭据（例如密码、密钥）的管理是科学家和信息安全官员的负担。凭据问题（例如，过期、特权不匹配）导致工作流无法获取所需的输入数据或存储有价值的科学结果，从而通过要求科学家诊断问题、重新运行其计算，并等待更长的时间，等待他们的结果。在本文中，我们介绍了科学令牌，开源软件，以帮助科学家更可靠和安全地管理他们的安全凭据。我们描述了科学令牌系统的架构、设计和实现，解决激光干涉仪引力波观测台（ligo）科学协作和大型天气测量望远镜（lsst）项目中的用例。我们还介绍了与支持分布式科学计算的广泛使用的软件的集成，包括 htcondor、cvmfs 和 xrootd。本令牌使用 ietf 标准 oauth 令牌对远程科学数据进行基于功能的安全访问。访问令牌传达工作流所需的特定授权，而不是通用身份验证模拟凭据，以解决在分布式基础结构（包括 nsf 资源）上运行的科学工作流的风险（例如，ligo 数据网格、开放科学网格、xsede）和公共云（例

如,亚马逊 web 服务、google 云、microsoft azure)。通过提高科学工作流的互操作性和安全性,科学令牌 1) 支持将分布式计算用于需要更大数据保护的科学领域, 2) 通过减少更广泛的分布式计算资源,实现了对更广泛的分布式计算资源的使用。远程系统上的凭据滥用的风险。少

2018 年 7 月 12 日提交;最初宣布 2018 年 7 月。

#### 47. 深度学习对情感分类的不同词语嵌入研究综述

作者:德巴德里·达塔

摘要: 网络上充满了文本内容,自然语言处理是机器学习中最重要领域之一。但是,当数据巨大时,简单的机器学习算法就无法处理它,而当基于神经网络的深度学习发挥作用时。然而,由于神经网络无法处理原始文本,我们必须通过一些不同的词嵌入策略来改变它们。本文演示了在 amazon 评论数据集上实现的那些独特的词嵌入策略,它有两种需要分类的情绪:基于众多客户评论的快乐和不快乐。此外,我们证明了在准确性的区别与关于嵌入哪个词应用的话语什么时候。少

2018 年 7 月 5 日提交;最初宣布 2018 年 7 月。

#### 48. 空间网: 遥感数据集与挑战系列

作者:adam van etten, dave lindenbaum , todd m. bacastow

**摘要:** 在世界许多地区,特别是在自然灾害等动态情况下,及时更新至关重要的情况下,基础制图仍然是一项挑战。更新地图目前是一个高度手动的过程,需要大量的人工贴标器来创建功能或严格验证自动输出。我们建议,经常对地球成像卫星星座进行重新审视,可以加快现有的努力,在与先进的机器学习技术相结合的情况下,快速更新基础地图。因此,spacenet 合作伙伴(cosmiq works、辐射解决方案和 nvidia)在 amazon web services (aws) 上发布了大量标记的卫星图像,称为 spacenet。spacenet 合作伙伴还发起了一系列公共奖项竞赛,以鼓励改进遥感机器学习算法。前两次比赛的重点是自动建筑足迹提取,最近的挑战集中在路网提取上。在本文中,我们讨论了 spacenet 图像、标签、评估指标、迄今为止的奖励挑战结果以及 spacenet 挑战系列的未来计划。少

**2018 年 9 月 25 日提交;**v1 于 2018 年 7 月 3 日提交;**最初宣布** 2018 年 7 月。

#### 49. 复位器: 自动智能 rest api 模糊

**作者:**[vaggelis atlidakis](#), [patrice godefroid](#), [marina polishchuk](#)

**摘要:** 随着亚马逊 web 服务和微软 azure 等强大云计算平台的出现,云服务最近爆发了。如今,大多数云服务都是通过 rest api 访问的,swagger 可以说是 rest api 最流行的界面描述语言。sw

款规范描述了如何通过其 rest api 访问云服务（例如，服务可以处理哪些请求以及可能会有哪些响应）。本文介绍了第一个自动智能 rest api 安全测试工具 rest ler。rest 程序分析 sw 款规范并生成测试，通过其 rest api 执行相应的云服务。每个测试都被定义为一系列请求和响应。rest ler 通过 (1) 推断 swagger 规范中声明的请求类型之间的依赖关系（例如，推断 "请求 b 不应在请求 a 之前执行" 来智能地生成测试，因为 b 将资源 id x 作为输入参数由 a) 和 (2) 通过分析在以前的测试执行过程中观察到的响应的动态反馈来生成新的测试（例如，了解 "请求 c 在请求序列 a 之后;b 被服务拒绝 "，因此在将来避免这种组合）。我们表明，这两种技术对于在修剪可能的请求序列的大搜索空间的同时，彻底执行正在测试的服务是必要的。我们还讨论了 rest ler 在测试 gitlab 方面的应用，这是一种大型流行的开源自托管 git 服务，以及发现的新错误。少

2018 年 6 月 25 日提交;最初宣布 2018 年 6 月。

## 50. 使用基于多模式注意的视频功能的端到端视听场景感知对话框

作者 :chiori hori, h 强化 alamri, jue wang, gordon wichern, takaaki hori, anoopcherian, tim k . marks, vincentcarartilier, raphael gontijo lopes, abhishek das, irfan essa, dhruv batra,devi parikh

摘要: 对话框系统需要了解动态视觉场景，以便与用户就其周围的对象和事件进行对话。可通过集成来自多个研究领域的最先进



技术（包括：端到端对话技术）来开发适用于实际应用的场景感知对话系统，这些技术使用从对话框数据；视觉问答（vqa）技术，利用学习的图像特征回答有关图像的问题；和视频描述技术，其中描述（描述）是利用多模式信息从视频中生成的。我们介绍了一个关于人类行为视频的新的数据集对话框。每个对话框都是类型化对话，由两个 **amazon** 机械土耳其人（amt）工作人员之间的 10 个问答（qa）对组成。我们总共收集了大约 9,000 个视频的对话框。利用这个新的音频视觉场景感知对话框（avsd）数据集，我们训练了一个端到端对话模型，该模型在有关视频的对话框中生成响应。我们的实验表明，使用为基于多模式注意的视频描述而开发的多模式功能可以提高生成的动态场景（视频）对话的质量。我们的数据集、模型代码和预培训模型将公开提供，以应对新的视频场景感知对话挑战。少

**2018 年 6 月 29 日提交**;**v1** 于 2018 年 6 月 21 日提交;**最初宣布** 2018 年 6 月。

## 51. 设计最佳二进制评级系统

**作者:**[nikhil garg](#), [ramesh johari](#)

**摘要:** 现代在线平台依靠有效的评级系统来了解项目。我们考虑在交易后收集  $\{\text{em 二进制反馈}\}$  的评级系统的优化设计。我们做了三个贡献。首先，我们将评级系统的性能正式化为它恢复项目

上的真实基础排名的速度（在很大程度上是偏差的意义上），同时考虑到这两个项目的基础匹配率和平台的偏好。其次，我们提供了一种有效的算法来计算产生最高这种性能的二进制反馈系统。最后，我们展示了如何使用这一理论视角来经验性地设计一个可实现的、近似最优的评级系统，并使用在 **amazon** 机械土耳其人身上收集的真实世界实验数据验证我们的方法。少

**2018 年 10 月 27 日提交**;v1 于 2018 年 6 月 18 日提交;**最初宣布** 2018 年 6 月。

## 52. 从协作筛选的评论中学习分布式表示

**作者** :[amjad almahairi](#), [kyle kastner](#), [kyunghyuncho](#), [aaron courville](#)

**摘要**: 最近的工作表明，基于混合筛选器的推荐系统可以通过纳入附带信息（如自然语言评论）来改进，以此作为使派生产品表示正规化的一种方式。在这种方法成功的推动下，我们引入了两种不同的评审模型，并研究了它们对协同过滤性能的影响。虽然以前最先进的方法是基于潜在的 dirichlet 分配 (lda) 模型的审查，我们探索的模型是基于神经网络：一个小词产品专家模型和一个经常性的神经网络。我们证明，专家产品模型提供了更大的灵活性，使其能够在 **amazon** 评审数据集上实现最先进的性能，优于基于 lda 的方法。然而，有趣的是，递归神经网络提供的更大的建模能力似乎削弱了模型作为产品表示的规律性的能力。少

2018 年 6 月 18 日提交;最初宣布 2018 年 6 月。

### 53. 人群驱动的数据挖掘

作者:[柴承良](#),[朱凡](#),[李国良](#),[王建南](#),[郑玉典](#)

**摘要:** 许多数据挖掘任务不能通过自动匹配的过程（如情绪分析和图像分类）来完全解决。众包是利用人类认知能力来处理这些机器艰苦任务的有效方法。得益于公共众包平台，例如[亚马逊机械土耳其人](#)和[众木花](#)，我们可以轻松地让数十万普通工人（即人群）参与进来，以解决这些机器艰苦的任务。在本教程中，我们将调查和综合关于人群供电数据挖掘的广泛的现有研究。我们首先概述众包，然后总结众包挖掘中必须考虑的基本技术，包括质量控制、成本控制和延迟控制。接下来，我们将回顾人群驱动的数据挖掘操作，包括分类、集群、模式挖掘、使用人群的机器学习（包括深度学习、转移学习和半监督学习）和知识发现。最后，我们提供众包数据挖掘中新出现的挑战。少

2018 年 10 月 18 日提交;v1 于 2018 年 6 月 13 日提交;最初宣布 2018 年 6 月。

### 54. 领域适应的投影嵌入：不同领域中情绪分析的联合建模

作者:[jeremy barnes](#), [roman klinger](#), [sabine schulte im walde](#)

**摘要:** 由于受监督的分类器对领域的变化非常敏感, 因此对情绪分析的域适应具有挑战性。解决这一问题的两种最突出的方法是结构函数学习和自动编码器。然而, 它们要么需要很长的训练时间, 要么在分歧很大的领域受到很大影响。在跨语言情绪分析的最新进展的启发下, 我们提供了一个新的视角, 并将域适应问题作为嵌入投影任务。我们的模型将输入两个单域嵌入空间作为输入, 并学习将它们投影到一个双域空间, 这被联合优化为 (1) 跨域项目和 (2) 预测情绪。我们在 20 个源目标域对上进行了域适应实验, 用于情绪分类, 并在 11 个域对上报告了新的最先进的结果, 包括**亚马逊**域适应数据集和 2013 年和 2016 年 semeval 数据集。我们的分析表明, 我们的模型在相似的域上的性能与最先进的方法相当, 而在高度不同的域上的性能要好得多。我们的代码可 [https://github.com/jbarnesspain/domain\\_blse](https://github.com/jbarnesspain/domain_blse)

**2018 年 6 月 13 日提交;**v1 于 2018 年 6 月 12 日提交;**最初宣布** 2018 年 6 月。

## 55. 在仅语音设置中回答问题的韵律修改

**作者:** Aleksandr chuklin, aliaksei sendn, johnne trippas, enrique alfonseca, hanna silen, damiano spina

**摘要:** 数字助手的许多流行形式因素——如**亚马逊** echo、apple homepod 或 google home——使用户能够仅根据语音方式与助手进行对话。缺乏一个屏幕, 用户可以从中阅读文本或观看支持

的图像或视频, 这带来了独特的挑战。为了满足用户的信息需求, 我们认为需要针对此类仅语音交互优化答案的呈现。在本文中, 我们提出了一个任务, 评估的有效性, 韵律修改的目的是只语音回答问题。我们描述了一种众包设置, 在这种设置中, 我们根据用户的信息、自然度和识别答案关键部分的能力的多个维度评估这些修改的质量。此外, 我们还提出了一套简单的韵律修改, 利用各种声学线索突出答案的重要部分。少

2018 年 6 月 11 日提交;最初宣布 2018 年 6 月。

## 56. 通过多任务学习在观点文本数据中的可解释性建议

作者:王楠,王洪宁,易陵佳,岳贤

**摘要:** 通过解释自动生成的建议, 用户可以更明智、更准确地决定要使用哪些结果, 从而提高他们的满意度。在这项工作中, 我们开发了一个多任务学习解决方案, 用于可解释的推荐。通过联合张量分解, 集成了推荐} 的用户偏好建模和用于解释的 \texalitenoenoy 型内容建模这两个配套学习任务。因此, 该算法不仅预测用户对项目列表 (即推荐) 的偏好, 还预测用户在功能级别对特定项目 (即固执己见的文本解释) 的欣赏方式。在 amazon 和 yelp 评论的两个大型集合上进行的大量实验证实了我们的解决方案在推荐和解释任务方面的有效性, 与现有的几种

推荐算法相比。我们广泛的用户研究清楚地证明了我们的算法所产生的可解释的建议的实用价值。少

2018 年 6 月 9 日提交;最初宣布 2018 年 6 月。

## 57. 会话推荐系统

作者:孙月明,张毅

**摘要:** 一个个性化的谈话销售代理可能会有很大的商业潜力。亚马逊、易趣、jd、阿里巴巴等电子商务公司正在与用户一起试制此类代理商。然而,对这一课题的研究非常有限,现有的解决方案要么是基于单轮自组织搜索引擎,要么是传统的多轮对话系统。它们通常只在当前会话中使用用户输入,而忽略了用户的长期首选项。另一方面,众所周知,销售转化率可以在推荐系统的基础上大大提高,推荐系统根据过去的购买行为学习用户偏好,并优化以业务为导向的指标,如转化率或预期收入。在这项工作中,我们建议将对话系统和推荐系统中的研究集成到一个新颖的、统一的深层强化学习框架中,以构建一个个性化的会话推荐代理,优化每个会话的实用程序功能。少

2018 年 6 月 8 日提交;最初宣布 2018 年 6 月。

## 58. ermes: 跨语言情感分类中的表情符号表示学习

作者:陈振鹏,沈生,胡志牛,宣路,梅乔珠,刘玄哲

**摘要:** 大多数现有的情绪分析方法在很大程度上依赖于大量标记的数据, 这些数据通常涉及耗时且容易出错的手动批注。这些标记数据在不同语言之间的分布明显不平衡, 例如, 与其他语言的文本相比, 有更多的英文文本被标记, 这对跨语言情绪分析提出了重大挑战。有几种跨语言的表示学习技巧, 将从一种带有丰富标签示例的语言中学到的知识传递到标签少得多的另一种语言。然而, 由于机器翻译质量不完善, 连接两种语言的信号很少, 它们的性能通常是有限的。在本文中, 我们使用情感, 一种无处不在的情感语言, 作为跨语言情感分析的新桥梁。具体而言, 我们提出了一种半监督表示学习的方法, 通过情绪预测的任务, 以学习跨语言的表示文本, 可以捕获语义和情绪信息。然后利用所学的陈述来方便跨语言情绪分类。我们展示了我们在具有代表性的 **amazon** 审查数据集上的方法的有效性和效率, 该数据集涵盖三种语言和三种领域。少

**2018 年 6 月 7 日提交;最初宣布** 2018 年 6 月。

## 59. 在自治仓库中实施增强现实: 挑战与机遇

**作者:** [david puljiz](#), [gleb gorbachev](#), [björn hein](#)

**摘要:** 带有移动、机架机器人的自主仓库开始变得司空见惯, 亚马逊的 **kwa** 和 **swisslog** 的 **carrypik** 等系统已经在功能仓库中实施。然而, 这种仓库仍然需要人为干预, 以进行文物的挑选和维

护。在欧洲安全日志项目中,我们正在开发一种安全背心,用于移动机器人的安全关键测距和停止,一个改进的规划师,可以处理异构代理的大型车队,以及一个 ar 交互系统来导航和支持在这样的自动化环境中的人工。在这里,我们介绍 ar 交互模式,即导航、按 ar 选择和编写时开发的一般系统交互,以及整个系统概念和计划的未来工作。少

**2018 年 6 月 1 日提交;最初宣布** 2018 年 6 月。

## 60. 对情感无人机群的研究: 一项初步的人群来源研究

**作者:**[trong-huy d. nguyen](#), [kisper Grispino](#), [damian lyons](#)

**摘要:** 无人机群是由自主的无人飞行器组成的团队,作为一个集体实体。我们对人性化的无人机群感兴趣,使他们具备通过非语言动作在情感上影响人类用户的能力。在最近观察者通过观看舞蹈动作在情感上受到感动的启发下,我们调查了无人机群的运动是否以及如何能够对一般观众产生情感影响的问题。我们对亚马逊机械土耳其人的初步研究得出了一些有趣的发现,包括有希望的结果和挑战。少

**2018 年 5 月 31 日提交;最初宣布** 2018 年 6 月。

## 61. 技术报告: 关键价值商店中的乐观执行



作者: [duong nguyen](#), [aleksey charapko](#), [sandeep kulkarni](#), [murat demirbas](#)

**抽象:** cap 定理的局限性意味着, 如果在存在网络分区的情况下需要可用性, 则必须牺牲顺序一致性, 这是一种更自然的系统设计一致性模型。我们关注的问题是, 如果设计人员有一个算法可以正常工作, 顺序一致性, 但面临的基础键值存储提供了较弱(例如, 最终或因果)的一致性, 那么他应该做什么。我们提出了一种基于检测回滚的方法: 设计人员识别一个正确性谓词, 例如  $P$ , 并继续运行协议, 因为我们的系统监视  $P$ . 如果  $P$  被破坏(因为基础键值存储提供了较弱的一致性), 系统回滚并恢复计算在一个状态, 其中  $P$  持。我们使用在伏地魔键值存储上运行的实际图形应用程序来评估这种方法。我们在 **amazon** aws ec2 实例上进行部署的实验表明, 使用与监视的最终一致性可以提供 50—80% 与顺序一致性相比, 吞吐量增加。我们还表明, 监视本身的开销很低(通常小于 4%), 并且检测冲突的延迟很小。特别是, 超过 99.9% 在不到的情况下检测到的违规情况。50 在区域 aws 网络中的毫秒, 并在不到 5 在全球 aws 网络中的秒数。少

**2018年6月23日提交;** v1 于 2018年5月25日提交; **最初宣布** 2018年5月。

## 62. 历时语感进化的无监督检测

作者: [jean-françois delpech](#)

**摘要:** 大多数词都有几种意义和内涵, 由于语义的转变而随着时间的推移而演变, 因此, 多年来, 密切相关的词可能会获得不同甚至相反的含义。这种演变与语言和文化变化的研究非常相关, 但目前可用于历时语义分析的工具具有显著的内在局限性, 不适合实时分析。在本文中, 我们演示了随机向量技术的线性如何能够构建一致的单词嵌入 (或语义空间) 的时间序列, 然后可以在不损失任何时间段精度的情况下进行线性比较和组合历时语义转换。我们展示了这种方法如何产生多义词 (如 **amazon** 或 **apple**) 的时间轨迹, 使元素间出现语义漂移和性别偏见, 揭示了飓风或总统等稳定概念的不断变化的实例化。这种非常快速、线性的方法可以很容易地分布在许多处理器上, 以便在推特或脸书等社交媒体的实时流中遵循; 由此产生的、与时间相关的语义空间可以通过简单的加减法随意组合。少

**2018年5月30日提交;**v1于2018年5月29日提交;**最初宣布** 2018年5月。

### 63. 多项用编码回归: 通过数据编码实现最佳的散射光器缓解

**作者:**songzeli, [seyed mohammadreza mousavi kalan](#), [qian yu](#) , [mahdi soltanolkotabi](#), [a. salman avestimehr](#)

**摘要:** 我们考虑了在大型数据集上使用梯度下降训练最小二乘回归模型的问题。计算是在由主节点和多个工作节点组成的分布式系统上进行的。由于存在运行缓慢的机器 (散客) 以及各种通信

瓶颈, 此类分布式系统明显放缓。我们提出了 "多项式编码回归" (pcr), 可大幅降低散向者的影响, 减轻此类系统中的通信负担。 pcr 的关键思想是对存储在每个工作人员身上的部分数据进行编码, 以便可以将工作人员的计算视为在不同点上计算多项式。这允许主机通过插值此多项式来计算最终渐变。 pcr 显著降低了恢复阈值, 定义为主服务器在计算渐变之前必须等待的工作人员数。特别是, pcr 需要一个恢复阈值, 该阈值与每个工作人员可用的计算量/存储量成反比。相比之下, 最先进的散链缓解方案需要更高的恢复阈值, 而这只会在每个工作人员计算/存储负载中线性减少。我们证明了 pcr 的恢复阈值接近最小, 并且在一个因素 2 的最佳可能的方案。我们在 **amazon ec2** 上的实验表明, 与最先进的方案相比, pcr 使用自然产生的散光器将运行时数提高了  $1.50x \sim 2.36x$ , 用人工散链器提高了  $2.36x \sim 4.29x$ 。少

**2018 年 5 月 24 日提交;最初宣布 2018 年 5 月。**

#### **64. 图形挑战. org: 提高图形分析性能的条杆**

**作者:**[siddharth samsi](#), [vijay Gadepally](#), [michael hurley](#), [michael jones](#), [edwardkao](#), [sanjeev mohindra](#), [paul monticciolo](#), [albertreuther](#), [steven smith](#), [williamsong](#), [diane staheli](#), [jeremy kepner](#)

**摘要:** 图形分析系统的兴起产生了对测量和比较图形处理系统能力的新方法的需求。设计了 [mit/myansa/eiee](#) 图形挑战, 为激发研究和突出图形分析软件、硬件、算法和系统方面的创新提供

了一个明确界定的社区场所.graphchallenge.org 提供了广泛的预解析图形数据集、图形生成器、数学定义的图形算法、各种语言的示例串行实现以及用于衡量性能的特定指标。2017 年图表挑战组织收到来自 36 个组织的 111 位作者提交的 22 份材料。提交的材料突出了硬件、软件、算法、系统和可视化方面的图形分析创新。这些提交材料产生了许多可比的性能计量,可用于评估该领域的最新技术状况。有许多提交材料实施了三角计数挑战,并导致超过 350 个不同的测量。对这些提交材料的分析表明,它们的执行时间是图形中边缘数的一个强大函数, $n_e$ ,并且通常与  $n^4$  个 / 3 个  $e$  对于大的值  $n_e$ 。结合提交的模型,可以了解图形分析的最新技术现状,这通常是 108 每秒处理的图形的边。108 边缘。这些结果是 30 比许多图形分析人员通常使用的串行实现快倍,并强调了向更广泛的社区提供这些性能优势的重要性。图形挑战提供了当前图形分析系统的清晰图景,并强调需要新的创新,以实现在非常大的图形上的高性能。少

**2018 年 5 月 22 日提交;最初宣布** 2018 年 5 月。

## 65. 高效 8 位低精度的卷积神经网络推理

**作者:**琼, 沈海豪, 张国明, 刘晓丽, 李善健, 葛进, 尼哈里卡·马赫赫瓦里, 埃法里德·福门科, 伊登·塞加尔

**摘要:** 神经网络的高吞吐量和低延迟推理是深度学习应用程序部署的关键。本文介绍了 intel 议事 f 的有效推理技术, 该技术是英特尔第一个支持英特尔至强卷积神经网络的 8 位低精度推理和模型优化技术的英特尔优化深度学习框架。可扩展处理器。8 位优化模型通过 fp32 模型的校准过程自动生成, 无需微调或再培训。结果表明, resnet-50、感知 v3 和 ssd 的推理吞吐量和延迟分别提高了 1.38x-2.9 x 和 1.35x-3x, 可忽略英特尔 caffe fp32 基线的精度损失, 以及 bvlc\_caffe 的 56x-75x 和 26x-37x。所有这些技术都是在 intel 议事 f github 上公开来源的, 提供该工件是为了在 amazon aws 云上重现结果。少

2018 年 5 月 4 日提交;最初宣布 2018 年 5 月。

## 66. 礼服: 融合的数据密集型计算平台的动态资源预留方案

**作者:**应茂,维多利亚·格林,王嘉音,熊浩义,郭志山

**摘要:** 在过去几年中, 我们设想越来越多的企业开始通过大数据分析 (如亚马逊推荐和 google 广告) 进行驱动。在后端, 企业由大数据处理平台提供动力, 以快速提取信息并做出决策。这些平台在计算集群之上运行, 利用调度算法来分配资源。由于资源有限 (如 cpu 和内存) 以及大量用户需求, 高效的调度程序对系统性能至关重要。但是, 除了来自客户端的请求和系统的当前状态外, 它对正在运行的作业的执行长度和传入作业的资源需求的了

解有限,这使得分配资源成为一项具有挑战性的任务。如果大部分资源被长时间运行的作业占用,则其他作业将不得不一直等待,直到释放它们。本文提出了一种新的调度策略,名为 dress,特别旨在优化具有各种需求的作业之间的分配。具体来说,它根据作业的请求将其分为两类,为每个类别保留一部分资源,并通过监视挂起的请求和估计正在运行的作业的发布模式来动态调整保留比率。结果表明,解和同时,解和保持了稳定的整体系统性能,显著缩短了一个类别的完成时间,最高可达 7.6.1%。少

**2018 年 5 月 21 日提交;最初宣布** 2018 年 5 月。

## 67.利用上下文完整性发现智能家庭物联网隐私规范

**作者:**noah apthorpe, yan shvartzshnaider, arunesh mathur, dillon reisman, nick feamster

**摘要:**面向消费者 "智能" 家居的物联网 (iot) 设备的激增引发了人们对用户隐私的担忧。我们提出了一种基于上下文完整性 (ci) 隐私框架的调查方法,该方法可以快速、高效地发现大规模的隐私规范。我们应用该方法在智能家居环境中发现隐私规范,在亚马逊机械土耳其人身上调查了 1731 美国成年人。在不到 6 小时的时间内,我们以 2,800 美元和不到 6 小时的时间内测量了 3,840 个信息流的可接受性,这些信息流代表了智能家居设备在各种条件下向第一和第三方收件人发送消费者信息的组合空间。我

们的研究结果为物联网设备制造商提供了可操作的建议, 包括设计最佳实践和采用我们的方法进行进一步研究的说明。少

2018 年 5 月 15 日提交;最初宣布 2018 年 5 月。

## 68. 异构群集的 sgx 感知容器业务流程

作者:sébastien vaucher, rafael pires, pascal felber, marcelo pasin, valerio schiavoni, christof fetzer

**摘要:** 容器正在成为在云中打包和部署应用程序和微服务的实际标准。一些云提供商(例如**亚马逊**、谷歌、微软)开始通过在其云产品中集成容器编排工具, 为其基础架构提供本机支持。同时, 集装箱向应用程序提供的安全保障仍然值得怀疑。客户仍然需要在数据和代码完整性方面信任他们的云提供商。英特尔最近在大众市场引入了软件保护扩展 (sgx), 这为开发人员提供了一种替代方案, 开发人员现在可以在硬件安全的环境中执行代码, 而无需信任云提供商。本文提供了有关 sgx 在行业标准的容器协调器 kubernetes 内部的支持的见解。我们在支持 sgx 的容器执行的整个堆栈中提供我们的贡献。我们提供有关调度程序及其监视框架的体系结构、基础操作系统支持和所需内核驱动程序扩展的详细信息。我们使用真实世界的 google borg 跟踪来评估我们在私有集群上的完整实现。我们的实验突出了在云中部署支持 sgx 的微服务时将遇到的性能权衡。少



2018 年 7 月 27 日提交;v1 于 2018 年 5 月 15 日提交;最初宣布 2018 年 5 月。

## 69. 一种经济有效的偏好启发式和聚合框架

作者:赵志明,李浩明,王俊明,杰弗里·凯哈特,尼古拉斯·马泰伊,许丽龙

**摘要:** 我们提出了一个具有成本效益的框架, 优先激发和聚集下的 Plackett–Luce 模型的特点。给定预算, 我们的框架迭代计算最具成本效益的启发式问题, 以帮助代理做出更好的组决策。我们通过在亚马逊机械土耳其人身上的实验来说明框架的可行性, 我们用实验来估计回答不同类型的启发式问题的成本。我们比较我们的框架的预测准确性时, 采用各种信息标准来评估预期的信息收益从一个问题。我们的实验表明, 精心设计的信息标准比随机提问要有效得多, 即使用较少的查询获得正确答案。少

2018 年 7 月 7 日提交;v1 于 2018 年 5 月 14 日提交;最初宣布 2018 年 5 月。

## 70. 无配对情感到情感的翻译: 一种循环强化学习的方法

作者:徐晶晶,孙秀蓝,曾琪,任宣城,张晓东,王厚峰,李文杰

**摘要:** 情感到情感 "翻译" 的目的是在保留句子内容的同时, 改变句子的基本情感。主要挑战是缺乏平行数据。为了解决这个问



题, 我们提出了一种循环强化学习方法, 通过中和模块和情绪化模块之间的协作, 实现了对未配对数据的训练。我们在 yelp 和 amazon 这两个评审数据集上评估我们的方法。实验结果表明, 我们的方法明显优于最先进的系统。特别是该方法大大提高了内容的保存性能。两个数据集的 BLEU 评分分别从 1.64 提高到 22.46 分和 0.56 分提高到 14.06 分。少

**2018 年 8 月 5 日提交;**v1 于 2018 年 5 月 14 日提交;**最初宣布** 2018 年 5 月。

## 71. 以消费者为中心的数据控制、跟踪和透明度-----立场文件

**作 者 :** james tapsell, raja naeem akram, konstantinos markantonakis

**摘要:** 与用户的活动、偏好和服务相关的个人数据被认为不仅是谷歌、亚马逊和苹果等众多以技术为导向的公司的宝贵商品, 也是像谷歌、**亚马逊**和苹果这样的更传统的公司的宝贵商品。旅游运输、银行、娱乐和营销行业。这导致了更有针对性的个人个性化服务-----在大多数情况下, 他们的财务费用最低。用户授权公司收集其个人数据以接收更多的个性化/目标/上下文感知服务和无忧活动(针对用户)的操作现实得到了广泛的部署。显然, 所收集数据的安全性、完整性和可访问性至关重要。在物联网 (iot)、自主车辆和无缝旅行的时代, 这些特征变得越来越根深蒂固。在本立场文件中, 我们将探讨用户和组织在处理个人身份信息 (pii)

方面所面临的挑战。此外,我们还进一步阐述了《一般数据保护条例》(gdpr)对 pii 管理的具体影响。随后,我们将讨论扩展到未来的技术,特别是物联网和集成运输系统,以获得更好的客户体验,以及它们在数据治理和 pii 管理方面的扩展。最后,我们提出了一个平衡用户隐私和数据控制的框架,以及一个组织的目标,即使用 "收集的用户数据" 向客户提供高质量、有针对性和高效的服务。此框架被称为 "面向消费者的数据控制 \& amp; 审核性" (codca), 并定义了适合隐私问题和法律监管框架的技术。少

**2018 年 5 月 12 日提交;最初宣布** 2018 年 5 月。

## 72. 将中本共识扩大到每秒数千笔交易

**作者:**李晨兴,李佩伦,周东,徐伟,范龙,姚德凡

**摘要:** 本文介绍了 conf 加斯,这是一个快速、可扩展和分散的区块链系统,它优化处理并发块,而不会丢弃任何作为分叉的块。comp 合计共识协议将块之间的关系表示为直接无环图,并在块的总顺序上达成共识。然后,从块顺序中确定地将交易总额订单派生为区块链分类帐。我们评估了 **amazon** ec2 集群上的 con 流量,并使用了多达 20k 的完整节点。conf 通在 4.5–7.4 分钟内确认事务,实现了 5.76bph 的事务吞吐量。对于典型的比特币交易,吞吐量相当于每秒 6400 笔交易。我们的研究结果还表明,在

运行 conf 通时, 协商一致协议不再是吞吐量瓶颈。相反, 瓶颈在于单个节点的处理能力。少

2018 年 8 月 31 日提交;v1 于 2018 年 5 月 10 日提交;最初宣布 2018 年 5 月。

### 73.slugnerds: 开放领域对话系统的命名实体识别工具

作者: [kevin k. bowden](#), [jiquiwu](#), [shereen oraby](#), [amita misra](#), [marilyn walker](#)

**摘要:** 在对话系统中, 命名实体识别 (ner) 和命名实体链接 (nel) 的任务是理解用户意图的重要预处理步骤, 特别是在开放域交互中, 我们不能依赖特定于域的推理。ucsc 作为 2017 年亚马逊亚历克莎奖竞赛中获得资助的团队之一所做的努力, 产生了 slugbot, 一个开放的领域的社交机器人, 旨在进行随意的对话。在构建 slugbot 时, 我们发现了一些与 ner 和 nel 特别相关的挑战, 例如 ne 标签的粗粒度太粗, 或者实体类型没有链接到有用的本体。此外, 我们发现, 传统方法在我们的背景下效果并不好: 即使是设计用于在推特或其他社交媒体数据上运作的系统, 在对话系统中也不能很好地发挥作用。在本文中, 我们介绍了 slugbot 的命名实体识别对话系统 (slugnerds), 这是一个 ner 和 nel 工具, 经过优化, 以解决这些问题。作为这项工作的一部分, 我们描述了两个新的资源: sl 空 enttdb 和架构执行器。我们相信这些资源将对研究界有用。少

2018 年 5 月 9 日提交;最初宣布 2018 年 5 月。

#### 74. 基于神经自动编码器决策林的意见欺诈检测

作者:董曼青,姚丽娜, 王贤志,布阿莱姆·贝纳塔拉,黄朝兰,宁晓东

**摘要:** 在线评论在影响买家的日常购买决定方面发挥着重要作用。然而,虚假的和毫无意义的评论,不能反映用户的真实购买体验和意见,广泛存在于网络上,并对用户做出正确的选择构成巨大挑战。因此,最好建立一个公平的模型,通过区分垃圾邮件审查来评估产品质量。我们提出了一个端到端可培训的统一模型,以利用自动编码器和随机林的吸引人的属性。实现了一个随机决策树模型,以指导全局参数学习过程。在大型亚马逊审查数据集上进行了大量实验。该模型的性能始终优于一系列比较方法。少

2018 年 5 月 9 日提交;最初宣布 2018 年 5 月。

#### 75. 主题: 生成特定于上下文的假餐厅评论

作者:mika juuti, bo sun, tatsuya mori, n. as 藤 an

**摘要:** 自动生成的假餐厅评论是对在线评论系统的威胁。最近的研究表明,用户很难检测到机器产生的假评论隐藏在真正的餐厅评论。这项工作中使用的方法(char-lstm)有一个缺点:它难以停留在上下文中,即当它为特定目标实体生成审阅时,所产生的评审可能包含与目标无关的短语,从而增加其可探测性。在这项工

作中，我们提出并评估了一种基于神经机器翻译 (nmt) 的更复杂的技术，我们可以用它来生成停留在主题上的评论。我们使用亚马逊机械土耳其人的母语为英语的人测试我们的技术的多种变体。我们证明，最好的变种产生的评论几乎具有最佳的不可探测性（类平均 f-分数 47%）。我们对持怀疑态度的用户进行了用户研究，并表明，与具有统计意义的最先进（平均规避 3.2/4 对 1.5/4）的最先进（平均规避 3.2/4 对 1.5/4）相比，我们的方法更频繁地避免检测，其水平是 $\alpha = 1\%$ （第 4.3 节）。我们开发了非常有效的检测工具，并在对这些工具进行分类时达到了 97% 的平均 f 分。虽然假评论是非常有效的愚弄人，有效的自动检测仍然是可行的。少

2018 年 6 月 28 日提交;v1 于 2018 年 5 月 7 日提交;最初宣布 2018 年 5 月。

## 76. 可持续云计算：基础和未来方向

作者: [Rajkumar buyya](#), [sukhpal singh gill](#)

**摘要:** 微软、谷歌、脸书和亚马逊等主要云提供商严重依赖数据中心来支持不断增长的对其计算和应用服务的需求。然而，运行如此庞大的基础设施所需的财务和碳足迹成本对云服务的可持续性产生了负面影响。现有的大部分工作主要集中在最大限度地减少服务器的能耗上。在本文中，我们为所有资源（包括服务器、网络、

存储、冷却系统) 的整体管理设计了一个概念模型和实用设计指南, 以提高云数据中心 (cdc) 的能效并减少碳足迹。).此外, 我们还讨论了可持续云计算的能源和可靠性之间的相互关联的关系, 并强调了相关的研究问题。最后, 我们提出了该领域未来的研究方向, 并为进一步的实际发展奠定了基础。少

2018 年 5 月 1 日提交;最初宣布 2018 年 5 月。

## 77. 了解和减轻亚马逊亚历克莎和谷歌首页语音控制的第三方技能的安全风险

作者:张楠,米香航, 玄峰, 王晓峰,袁天, 钱峰

**摘要:** 虚拟个人助理 (vpa) (如 **amazon** 亚历克莎和 **google** 助手) 今天主要依靠语音通道与用户沟通, 但众所周知, 用户是脆弱的, 缺乏适当的身份验证。vpa 技能市场的快速增长开辟了一条新的攻击途径, 有可能允许远程对手发布攻击技能, 通过热门的物联网设备 (如 **amazon** echo 和 **google** home) 攻击大量 vpa 用户。在本文中, 我们报告了一项研究, 得出这种远程、大规模攻击确实是现实的。更具体地说, 我们实施了两种新的攻击: 语音蹲, 其中对手利用技能的调用方式 (例如, "开放资本一"), 使用具有类似发音的名称 (例如, "资本赢得") 或转述名称 (例如, "资本一请 ") 劫持语音命令, 意味着不同的技能, 和语音伪装, 其中恶意技能模拟 vpa 服务或合法技能窃取用户的数据或窃听她的对话。

这些攻击针对的是 vpa 的工作方式或用户对其功能的误解, 并被我们在 **amazonecho** 和 **google home** 上的实验 (包括用户研究和实际部署) 发现构成了现实的威胁。我们的发现的重要性已经得到了**亚马逊**和**谷歌**的认可, 我们构建的新检测系统在亚历克莎和谷歌市场上发现的高风险技能进一步证明了这一点。我们进一步开发了自动检测这些攻击的技术, 这些技术已经抓住了可能构成此类威胁的现实世界技能。少

**2018 年 6 月 29 日提交;v1** 于 2018 年 5 月 3 日提交;**最初宣布** 2018 年 5 月。

## 78. 一种分层的端到端模型, 用于共同改进文本摘要和情感分类

**作者:**[马树明](#),[孙旭](#),[林俊阳](#),[任宣城](#)

**摘要:** 文本摘要和情感分类都是为了捕捉文本的主要思想, 但在不同的层面上。文本摘要是在几个句子中描述文本, 而情感分类可以看作是一种特殊的总结类型, 它将文本 "总结" 为一种更加抽象的方式, 即情感类。基于这一思想, 我们提出了一个分层的端到端模型, 用于文本摘要和情感分类的联合学习, 将情感分类标签作为文本摘要输出的进一步 "摘要"。因此, 将情感分类层放在文本摘要层上, 导出了层次结构。**亚马逊**在线评论数据集的实验结果表明, 我们的模型在抽象摘要和情绪分类方面都比强基线系统具有更好的性能。少

**2018 年 5 月 29 日提交;v1 于 2018 年 5 月 2 日提交;最初宣布** 2018 年 5 月。

## 79. 但谁来保护版主呢？众包图像适度的案例

**作者:**[brandon dang](#), [martin j. riedl](#),[matthew lease](#)

**摘要:** 尽管已经开发了检测系统来识别色情和暴力等淫秽内容，但人工智能根本还不足以完全实现这一任务的自动化。由于需要人工验证，社交媒体公司可能会聘请内部审查人员、与第三方的专业工人签订合同，或将其外包给在线劳动力市场，以进行商业内容适度。这些内容主持人往往完全接触到极端的内容，可能会遭受持久的心理和情感伤害。在这项工作中，我们的目标是通过调查以下问题来缓解这一问题：我们如何向人类审阅者揭示最小的信息量，以便仍然能够正确识别令人反感的图像？我们设计并进行实验，其中模糊的图形和非图形图像被亚马逊机械土耳其 (amt) 上的人类版主过滤。我们观察模糊如何影响在图像分类准确性、界面可用性和员工情感健康方面的适度体验。少

**2018 年 6 月 9 日提交;v1 于 2018 年 4 月 29 日提交;最初宣布** 2018 年 4 月。

## 80. 发声板：以用户为中心、内容驱动的交流聊天机器人

**作者:**[郝芳](#),[郝成正](#), [maarten sap](#), [elizabeth clark](#), [ari holtzman](#), [yejin choi](#), [noah a. smith](#), [mari ostendorf](#)



**摘要:** 我们介绍了探测委员会, 这是一个获得 2017 年**亚马逊**亚历克莎奖的社交聊天机器人。系统体系结构由多个组件组成, 包括口语处理、对话管理、语言生成和内容管理, 重点是以用户为中心和内容驱动的设计。我们还分享从基于 160, 000 次与现实世界用户对话的大规模在线日志中获得的见解。少

**2018 年 4 月 26 日提交;最初宣布** 2018 年 4 月。

## 81. 基于对抗性剩余变换网络的无监督域适应

**作者:**[蔡冠宇](#),[王玉琴](#),[周梦楚](#),[何良华](#)

**文摘:** 领域适应在缺乏标签的学习问题中得到了广泛的应用。最近的研究表明, 深度对抗域适应模型可以在性能上进行显著的改进, 其中包括对称和非对称体系结构。然而, 前者的泛化能力较差, 而后者则很难培养。本文提出了一种新的对抗域适应方法--对抗性剩余变换网络 (artn), 以提高泛化能力, 直接将源特征转化为目标特征的空间。在该模型中, 利用剩余连接共享特征, 重建对抗性损失, 使模型更加通用, 更容易训练。此外, 在损失函数中加入正则化, 以缓解消失梯度问题, 使训练过程稳定。基于**亚马逊**审查数据集、数字数据集和 office 31 图像数据集的一系列实验结果表明, 所提出的 artn 方法大大优于最先进的方法。少

**2018 年 4 月 25 日提交;最初宣布** 2018 年 4 月。

## 82. 年龄改变对语音伪装有效性的感知评价

作者: [rosa gonzález Hautamäki](#), [anssi kanervisto](#), [ville Hautamäki](#), [tomi kinnunen](#)

**摘要:** ...。我们进行了一个感性的测试, 以研究伪装语音样本的质量。听力测试是在当地进行的, 并得到了**亚马逊机械土耳其人** (mt) 的帮助。共有 91 名听众参加了测试, 并被指示估计发言者的时间顺序和预期年龄。那是..。更多

**2018 年 5 月 28 日提交;**v1 于 2018 年 4 月 24 日提交;**最初宣布** 2018 年 4 月。

## 83. 云托管虚拟网络功能的价格与性能: 分析与未来挑战

作者: [nadir ghrada](#), [mohamed faten zhani](#), [yehia elkhatib](#)

**摘要:** 网络功能虚拟化 (nfv) 的概念近年来被引入到一种新的范式中。nfv 提供了许多好处, 包括显著提高可维护性和减少部署开销。已经做了几项工作来优化虚拟网络功能 (vnf) 的部署 (也称为嵌入)。但是, 到目前为止, 还没有任何工作研究如何优化给定 vnf 的云实例选择及其特定要求。在本文中, 我们评估了 vnf 在不同的 **amazon** ec2 云实例上嵌入时的性能。具体来说, 我们从到达数据包速率、资源利用率和数据包丢失等方面评估三个 vnf (防火墙、ids 和 nat)。我们的结果表明, 不同实例类型的性能各不相同, 这背离了云实例 "您得到您所支付的费用" 的直觉。

我们还发现 cpu 是经过测试的 vnf 的关键资源, 尽管它们的峰值数据包处理能力差别很大。最后, 在所得结果的基础上, 确定了与 vnf 实例选择和服务链配置有关的关键研究挑战。少

2018 年 4 月 23 日提交;最初宣布 2018 年 4 月。

#### 84. 在本地衰减, 在全球范围内取胜: 基于衰减的在线随机匹配框架

作者 :[brian brubach](#), [karthik abinav sankararaman](#), [aravind srinivasan](#), [pan xu](#)

**文摘:** 近年来, 由于在线广告、在线广告、坐骑共享等领域的大量应用, 在线匹配问题受到了广泛的关注。他们中的许多人通过将随机性包括在到达过程和匹配过程中, 捕捉到了现实世界中的不确定性。banal 等人提出的在线随机匹配与超时问题 (算法, 2012 年) 模型匹配市场 (例如, e-bay, **amazon**)。买家来自一个独立的和相同分布 (i. d.) 已知的分布买家配置文件, 可以显示一个项目的列表一次。每个买家都有购买每个项目的一定概率, 并对他们可以显示的物品数量有一定的限制 (超时)。bansal 等人 (算法, 2012 年) 给出了一个 0.12 竞争算法, adamczyk 等人 (欧空局, 2015 年) 将该算法改进为 0.12。我们提出了一个在线衰减框架, 该框架使用离线随机匹配作为黑匣子的算法。在上限方面, 我们表明, 这个框架, 加上一个从 bansal 等人 (算法, 2012 年) 改编的黑匣子, 产生了一个在线算法, 该算法的比率几乎翻倍, 达到 0.46。在下界方面, 我们证明, 对于这个问题, 没有任何算法能达

到比 0.632 更好的比率。该框架具有很高的进一步改进潜力，因为新的离线随机匹配算法可以直接提高在线问题的比率。我们的在线框架也有可能进行各种扩展。例如，我们引入了一个自然泛化：在线随机匹配与双面超时，其中在线和离线顶点都有超时。我们的框架为这一问题提供了第一个算法，实现了 0.30 的比率。我们再次使用 adamczyk 等人的算法 (esa, 2015) 作为黑匣子，并将其插入我们的框架。少

**2018 年 4 月 21 日提交;最初宣布** 2018 年 4 月。

#### 85. alquist: 亚历克莎奖社会机器人

**作者:**janpichl, petr marek, jakub konrád, martin matulík, hoang long nguyen, janš edivüu

**文摘:** 本文介绍了一个新的开放领域对话系统 alquist 开发的亚历克协会奖竞争的亚马逊 echo 系列产品的一部分。阿尔奎特对话系统旨在就热门话题进行连贯和引人入胜的对话。我们提出了一个混合系统，结合了几种机器学习和基于规则的方法。我们讨论和描述了 alquist 管道、数据采集和处理、对话管理器、nlg、知识聚合和子对话的层次结构。我们给出了一些实验结果。少

**2018 年 4 月 18 日提交;最初宣布** 2018 年 4 月。

#### 86. 删除、检索、生成: 一种简单的情绪和风格转换方法

作者:李俊森,贾罗宾,何河,梁珀西

**摘要:** 我们考虑文本属性传输的任务: 转换句子以更改特定属性 (例如, 情绪), 同时保留其与属性无关的内容 (例如, 将 "屏幕只是合适的大小" 更改为 "屏幕太小")。我们的训练数据只包括带有属性标记的句子 (例如, 正面或否定), 而不包括仅在属性上不同的句子对, 因此我们必须学会在无人监督的情况下将属性与属性无关的内容区分开来方式。以前使用敌对方法的工作一直难以产生高质量的产出。在本文中, 我们提出了更简单的方法, 其动机是观察到文本属性通常由独特的短语 (例如, "太小") 标记。我们最强的方法是通过删除与句子原始属性值关联的短语、检索与目标属性关联的新短语以及使用神经模型将这些短语流畅地组合到最终输出中来提取内容词。在人的评价方面, 我们最好的方法是在比以前最好的系统多 22% 的投入上产生语法和适当的反应, 平均在三个属性转移数据集上: 改变对 yelp 的评论的看法, 改变对亚马逊, 并改变图像标题是更浪漫或幽默。少

2018 年 4 月 17 日提交;最初宣布 2018 年 4 月。

## 87. 支持条件属性和自定义属性的可扩展权限管理系统

作者:刘白玉, abhinav palia, shan-ho yang

**摘要:** 随着不同企业管理多个身份、行动、设备、api 等的经典问题, 对具有灵活的基于属性的访问控制 ~ (abac) 机制的需求也

越来越大。为了填补这一空白, 提出了 abac 模型的几个变体, 如 \textit{amazon 的 aws iam}, 它使用 json 作为其基础存储数据结构, 并在常规 abac 上添加策略约束作为字段。但是, 这些系统仍然不提供自定义权限和对其执行各种操作 (如比较聚合) 的功能。在本文中, 我们引入了一个基于字符串的资源命名策略, 该策略支持资源访问的自定义和条件权限。此外, 我们还提出了我们系统的基本架构, 以及我们的命名方案, 使系统具有可扩展性、安全性、效率、灵活性和可定制性。最后, 给出了算法的概念证明, 以及实验设置和本工作的未来轨迹。少

2018 年 4 月 17 日提交;最初宣布 2018 年 4 月。

## 88. 众包中发布的定价: 奖金的力量

作者:suhoshin, jungseulok, hoyong choi, yung yi

**摘要:** 在亚马逊机械土耳其等实用的众包系统中, 由于其简单性, 张贴定价被广泛使用, 任务请求者会先验地发布定价规则, 工人可据此决定是否接受和执行任务, 并且往往支付根据他们的努力的质量。一个好的张贴定价的关键要素之一在于如何以更低的预算招聘更多的高质量工人, 为此考虑了以下两个方案: (一) 通过对用户的质量和成本进行分析来进行个性化定价, (二) 额外的奖金支付提供了更多的合格任务完成。尽管它们在众包定价方面有潜在的好处, 但人们对个性化和奖金支付实际上为请求者提供了

多少收益一直没有充分的探索。在本文中，我们研究了四种可能的组合张贴定价的定价与定价—没有个性化和奖金。我们的目标是分析量化这两个想法何时以及在多大程度上有助于请求者的效用。为此，我们首先推导出最优的个性化和通用定价方案，并分析其计算可追踪性。接下来，我们量化两个定价方案中有奖金支付和不支付奖金之间的效用差距。我们分析证明，在个性化定价中，奖金的影响可以忽略不计，而在共同定价中却至关重要。最后，我们研究了非相关价格的概念，它量化了个性化定价政策和共同定价政策之间的效用差距。这意味着，一个复杂的个性化定价与隐私问题可以被一个简单的共同定价与奖金取代。我们通过在amazon 机械土耳其公司进行的大量模拟和实际实验验证我们的分析结果，并提供了在设计定价政策时有用的其他含义。少

2018 年 4 月 9 日提交;最初宣布 2018 年 4 月。

## 89. 文本情感的自动分类

作者:emmanuel dufourq , bruce a. bassett

**摘要:** 在文本中识别情感的能力，被称为情绪分析，对成年人类来说是很自然的。但是，此任务不是计算机在默认情况下可以执行的任务。以自动化的算法方式识别情绪将是企业和研究在他们的搜索中的一个有用的能力，以了解消费者对其产品或服务的看法，并了解人类社会学。在这里，我们提出了两个新的遗传算法

(ga) 的任务自动文本情绪分析。ga 了解文本语料库中出现的单词是情感还是放大器词, 以及它们的相应大小。情绪的话, 如 '可怕', 为最终的情绪增添了线性。放大词, 通常是形容词, 像 '非常', 乘以下面的词的情绪。这增加、减少或否定了以下单词的情绪。全文的感想就是这些术语的总和。这种方法既增加了情感字典, 也增加了放大器字典, 可以用于其他目的, 并输入到其他机器学习算法中。我们报告在大型**亚马逊**数据集上进行的多项实验的结果。结果表明, 我们提出的方法能够超越几种公共和商业情绪分析算法。少

**2018 年 4 月 5 日提交;最初宣布** 2018 年 4 月。

## 90. iaas 中的计时通道: 如何识别和调查

**作者:**小福,杨瑞丽,杜晓江,罗斌

**摘要:** 最近, iaas (基础架构即服务) 云 (例如, **amazon ec2**) 已被许多组织广泛使用。但是, 一些 iaas 安全问题会对其用户造成严重威胁。一个典型的问题是计时通道。正如许多研究人员所证明的那样, 这种渠道可以是跨虚拟机的信息通道。由于其隐蔽性和无痕量, 传统的识别方法无法建立准确的分析模型, 也无法获得泄露的结果。我们从内存活动记录的角度研究了时序通道的基本行为, 总结了基础内存活动中计时通道的签名。提出了一种基于长期行为签名的识别方法。我们提出了一套完整的取证步骤,



包括证据提取、识别、记录储备和证据报告。我们研究了四个典型的定时通道, 实验表明, 即使在正常过程的干扰下, 也能检测和研究这些通道。少

2018 年 4 月 4 日提交;最初宣布 2018 年 4 月。

## 91. 通用数据库: 普及近似查询处理

作者: 永雄公园, barzan mozafari, joseph sorenson, junhao wang

**摘要:** 尽管在学术界进行了 25 年的研究, 但近似查询处理 (aqp) 几乎没有被工业采用。这种缓慢采用的主要原因之一是传统供应商不愿意对其传统代码库进行根本性的更改, 以及对实施标准功能的较新供应商 (例如 sql-hadoop 产品) 的关注。此外, 可用的几个 aqp 引擎都与特定平台相关联, 并要求用户完全放弃现有数据库——考虑到 aqp 技术的起步阶段, 这是一个不切实际的期望。因此, 我们认为需要一个通用的解决方案: 一个与数据库无关的近似引擎, 它将扩大这一新兴技术在各种平台上的影响范围。我们的建议称为 verfondb, 它使用的中间件体系结构不需要对后端数据库进行任何更改, 因此可以使用所有现成的引擎。在驱动程序级别运行时, verfoldb 会截获发送到数据库的分析查询, 并将其重写到另一个查询中, 如果由任何标准关系引擎执行, 该查询将产生足够的信息来计算近似答案。verfoldb 使用返回的结果集计算近似答案和错误估计, 然后将其传递给用户或应用程

序。但是, 由于无法访问查询执行层, 在通用性、正确性和效率方面带来了重大挑战。本文展示了 `verfilb` 如何克服这些挑战, 并提供高达 171X 加速 (18.45X 平均而言) 用于各种现有引擎, 如 `impala`、`spark sql` 和 **amazon redshift**, 而相对误差不到 2.6%。根据 `apache` 许可证, `verfisb` 是开源的。少

2018 年 6 月 15 日提交;v1 于 2018 年 4 月 2 日提交;最初宣布 2018 年 4 月。

## 92. 用于交互式分析的最先进的 `sql-hadoop` 系统的比较分析

作者:[ashish tapdiya](#), [daniel fabbri](#)

**摘要:** `hadoop` 正在成为企业的主要数据中心, `sql` 代表了数据分析的实际语言。这种结合导致了目前使用的各种 `sql-hadoop` 系统的开发。虽然各种 `sql-hadoop` 系统针对的是同一类分析工作负载, 但它们不同的体系结构、设计决策和实现会影响查询性能。在这项工作中, 我们使用 `web` 数据分析微基准和 **amazon ec2** 云平台上的 `tpc-h` 基准, 对四个最先进的 `sql on hadoop` 系统 (`impala`、`钻头`、`spark sql` 和 `phoenix`) 进行了比较分析。`tpc-h` 实验结果表明, 尽管 `impala` 在文本格式中的性能优于其他系统 (4.41 x–6.65 x), 但在实木复合地板格式中存在权衡, 每个系统在查询子集上表现最佳。对执行配置文件的全面分析扩展了性能结果, 以提供对性能变化、性能瓶颈和查询执行特征的洞察。少

2018 年 3 月 31 日提交;最初宣布 2018 年 4 月。

### 93. 熵控制非配置, 以提高独立学习者在匿名 marl 设置中的绩效

作者: [tanvi verma](#), [pradeep varakantham](#), [hoong chuin lau](#)

**摘要:** 随着 (供需) 顺序匹配系统的出现 (uber、lyft、起出租车; 食物的 ubereats、送货员等; amazon prime、lazada 等为日用品) 跨越许多在线和离线服务, 个人 (出租车司机, 送货男孩, 送货货车司机等) 赚更多的是在 "正确" 的时间在 "正确" 的地方。我们专注于学习技术, 以便在其他 "学习" 个人在场的情况下, 为个人提供指导 (关于在适当的时间提供指导)。独立之间的交互是匿名的, 即交互的结果 (争夺需求) 独立于代理的身份, 因此我们将这些称为匿名 marl 设置。现有的相关性研究是利用强化学习 (rl) 或多剂强化学习 (marl) 进行自主学习。聚合系统中的个人数量非常多, 个人有自己的私利 (收入最大化)。因此, 传统的 marl 方法要么不可扩展, 要么对共同目标或行动协调的假设不可行。本文着重研究了利用匿名性提高独立强化学习者的绩效, 特别是流行的深度 q 网络 (dqn) 和优势行为者批评 (a2c) 方法。具体而言, 我们使用代理密度分布熵来控制其他代理引入的非稳定性。我们展示了个人和所有代理的收入显著改善, 我们的学习者在为聚合系统和现实世界中的出租车数据集建立了一个通用的实验。少

2018 年 5 月 27 日提交;v1 于 2018 年 3 月 27 日提交;最初宣布 2018 年 3 月。

#### 94. 低复杂性分布式无线接入网切片: 算法与实验结果

作者 :salvatore d ' oro, francesco restuccia, tommaso melodia, sergio palazzo

**摘要:** 无线接入网 (ran) 切片是在 5g 网络中动态分配网络资源的一种有效方法。ran 切片的主要挑战之一是它显然是一个 np-hard 问题。为此, 我们设计了近乎最优的低复杂度分布 ran 切片算法。首先, 我们将切片问题建模为拥塞博弈, 并证明了这种博弈承认了一个独特的纳什均衡 (ne)。然后, 我们对 ne 的无政府制 (poa) 的价格进行了评价, 即 ne 与社会最优相比的效率, 并证明了它的上限为 "2"。接下来, 我们提出了两个完全分布的算法, 可以证明收敛到唯一的 ne, 而不显示来自切片租户的隐私敏感参数。此外, 我们还引入了无线资源的自适应定价机制, 以提高网络所有者的利润。我们通过模拟和部署在 **amazon** ec2 云上的实验测试平台来评估算法的性能, 这两者都基于 opencellid 项目中的基站的真实数据集。结果表明, 我们的算法收敛迅速, 达到了近乎最优的性能, 而我们的定价机制有效地提高了网络所有者的利润。少

2018 年 3 月 20 日提交;最初宣布 2018 年 3 月。

## 95. 基于神经网络的小型和大型认证验证实验

作者: [marjan hosseinia](#), [arjun mukherjee](#)

**摘要:** 对于作者身份验证问题的特殊情况, 我们提出了两种模型。任务是调查给定对的两个文档是否由同一作者编写。我们考虑了小型和大型数据集的作者身份验证问题。潜在的小规模问题有两个主要挑战: 第一, 文档的作者对我们来说是未知的, 因为以前没有书面样本。其次, 这两份文件很短(几百到几千字), 在体裁和/或话题上可能有很大的不同。为了解决这个问题, 我们建议变换编码器将一对的一个文档转换为另一个文档。此文档转换会生成一个损失, 该损失用作可识别的功能, 以验证对的作者是否相同。针对各种作者参与的大规模问题, 以及更多长度较大的例子, 提出了一种并行递归神经网络。比较了这两个文档的语言模型。我们评估我们对各种类型数据集的方法, 包括 pan 竞赛的作者身份标识数据集、amazon 评论和机器学习文章。实验表明, 与基线相比, 两种方法都取得了稳定和有竞争力的性能。少

2018 年 3 月 17 日提交;最初宣布 2018 年 3 月。

## 96. 折叠: 微型批次上的高效伪独立重量扰动

作者: [yemingwen](#), [paul vicol](#), [jimmy ba](#), [dustin tran](#), [roger grosse](#)

**摘要:** 随机神经网络权值用于多种背景, 包括正则化、贝叶斯神经网络、增强学习中的探索和进化策略。遗憾的是, 由于权重量大, 小型批次中的所有示例通常都具有相同的权重扰动, 从而限制了大型小型批次的方差减少效果。我们引入了一种有效的方法, 通过隐式采样每个示例的伪独立权重扰动来修饰小型批处理中的渐变。从经验上讲, 对于完全连接的网络、卷积网络和 mn, 触发器实现了理想的线性方差约简。我们发现在训练神经网络与乘法高斯摄动有显著的速度。我们证明了触发器在规范 lstm 方面是有效的, 并且优于以前的方法。flipout 还使我们能够对进化策略进行矢量化: 在我们的实验中, 使用现有方法的单个 gpu 可以处理与至少 40 个 cpu 内核相同的吞吐量, 相当于亚马逊 web 服务上 4 的成本降低。少

**2018 年 4 月 2 日提交;**v1 于 2018 年 3 月 12 日提交;**最初宣布** 2018 年 3 月。

## 97.利用众包数据进行深度主动学习-应用: 在亚历克莎学习意图

**作者:** jie yang, thomas drake , andreas damianou, yoelle maarek

**摘要:** 本文提出了一个通用的贝叶斯框架, 使任何深度学习模型都能积极地目标人群中学习。我们的框架继承了贝叶斯深度学习的最新进展, 并通过考虑有针对性的众包方法扩展了现有的工作, 在这种方法中, 具有未知专业知识的多个注释者贡献了一个

不受控制的数量（通常是有限的）注释。我们的框架利用注释中的低级别结构来学习单个注释器的专业知识，这有助于从嘈杂和稀疏的注释中推断真实的标签。它提供了一个统一的贝叶斯模型，同时推断真实的标签，训练深度学习模型，以达到最佳的学习效果。最后，我们的框架利用了预测过程中深度学习模型的不确定性以及注释者的估计专业知识，以最大限度地减少为优化培训深度学习模型而需要的注释和注释器的数量。我们使用合成数据集和真实世界数据集评估我们的意图分类框架在亚历克莎（**亚马逊**的个人助理）的有效性。实验表明，与最先进的方法相比，我们的框架能够准确地学习注释器的专业知识，推断真实的标签，并有效地减少模型培训中的注释量。我们进一步讨论了我们提出的框架在连接机器学习和众包方面对改进人与环系统的潜力。少

**2018 年 3 月 12 日提交;最初宣布** 2018 年 3 月。

## 98. 基于邻域的功能映射对冷启动用户的跨域推荐

**作者:**王兴华,彭兆辉,王森张,余志强, 傅文静, 洪晓光

**摘要:** 协同过滤 (cf) 是推荐系统中广泛采用的一种技术。传统的 cf 模型主要侧重于预测用户对单个域（如电影域或音乐域）中的项目的偏好。此类模型面临的一个主要挑战是数据稀疏问题，尤其是 cf 无法对完全没有评级的冷启动用户做出准确的预测。尽管提出跨域协作筛选 (cdcf) 可以在不同的域中有效地转移用户



的评级偏好,但现有的 cdcf 模型仍然难以解决目标域中的冷启动用户,因为极端的数据稀疏性。本文针对目标域中的冷启动用户,提出了跨域潜在特征映射 (cdlfm) 模型。首先,为了更好地描述稀疏域中的用户,我们考虑了用户在评级行为上的相似关系,并结合用户相似性 (mfus),结合三种相似性,提出了矩阵分解法。提出了措施。其次,为了实现跨域知识的转移,我们提出了一种基于邻域的梯度提升树方法来学习跨域用户潜在特征映射功能。对于每个冷启动用户,我们了解了基于那些在辅助域中与冷启动用户具有类似评级行为的链接用户的潜在功能对的功能映射功能。根据映射函数及其在辅助域中的潜在特性,可以预测目标域中冷启动用户的偏好。从 amazon 事务数据中提取的两个实际数据集的实验结果表明,我们提出的模型相对于其他最先进的方法具有优越性。少

2018 年 3 月 5 日提交;最初宣布 2018 年 3 月。

## 99. 强化学习在电子商务搜索引擎中排名: 形式化、分析和应用

作者:胡玉静,青达,曾安祥,杨宇,徐英辉

摘要: 在亚马逊、淘宝等电子商务平台中,在搜索会话中对项目进行排名是典型的多步骤决策问题。学习排名 (lir) 方法已被广泛应用于排名问题。但是,这类方法通常认为会话中不同的排名步骤是独立的,相反,这可能是高度相关的。为了更好地利用不同排名



步骤之间的相关性, 本文建议采用强化学习 (rl) 来学习最优排名策略, 最大限度地提高搜索会话中的预期累积奖励。首先, 我们正式定义了搜索会话马尔可夫决策过程 (ssmdp) 的概念, 以制定多步骤排序问题。其次, 分析了 ssmdp 的性质, 从理论上论证了最大限度地提高累积收益的必要性。最后, 提出了一种新的策略梯度算法来学习最优的排序策略, 该算法能够解决 ssmdp 的高报酬方差和不平衡的奖励分配问题。在仿真和淘宝搜索引擎上进行了实验。结果表明, 该算法的性能明显优于在线 ltr 方法, 在仿真和实际应用中, 总交易量分别增长了 40% 和 30% 以上。少

**2018 年 5 月 23 日提交;v1 于 2018 年 3 月 1 日提交;最初宣布 2018 年 3 月。**

## 100. 智能虚拟助手了解您的生活

**作者:**[钟贤基](#),[李桑金](#)

**摘要:** 在物联网世界中, 智能虚拟助理 (iva) 是一种流行的基于语音命令与用户交互的服务。为了获得最佳性能和高效的数据管理, 著名的 iva (如 **amazon** 亚历克莎和 **google** 助手) 通常基于云计算架构进行操作。在此过程中, 可以将大量包含用户语音活动历史记录以及详细说明的行为跟踪存储在 iva 生态系统中的远程服务器中。如果这些数据 (也称为 iva 云原生数据) 被攻击泄露, 恶意用户不仅可以收集 iva 服务的详细使用历史记录, 还

可以通过各种数据分析技术揭示其他与用户相关的信息。在本文中, 我们首先显示和分类类型的 iva 相关的数据, 可以收集从流行的 iva, **亚马逊**亚历克莎。然后, 我们分析了一个实验数据集, 涵盖三个月的亚历克莎服务, 并描述了用户的生活方式和生活模式的属性。我们的研究表明, 有可能发现关于个人信息的新见解, 如用户兴趣、iva 使用模式和睡眠、唤醒模式。本文给出的结果对 iva 供应商和用户的隐私威胁也有重要影响和威胁。少

2018 年 2 月 27 日提交;最初宣布 2018 年 3 月。

## 101. 百万富翁: 众包的一种有指导的方法

作者 :bo han, quming yao, yuangang pan, ivor w. tsang , xiaquixiao, 焦强, masashisugiyama

**摘要:** 现代机器学习正在迁移到复杂模型的时代, 这需要大量注释良好的数据。虽然众包是实现这一目标的一个有希望的工具, 但现有的众包方法几乎无法获得足够数量的高质量标签。本文以百万富翁游戏节目中的 "吸引人而暗示" 的回答策略为动力, 将提示引导的方法引入众包中, 以应对这一挑战。我们的方法鼓励工人在不确定问题时从提示中获得帮助。具体而言, 我们提出了一个混合阶段设置, 包括主阶段和提示阶段。当工人在主舞台上面临任何不确定的问题时, 他们被允许进入暗示阶段, 在做出任何答案之前先寻找暗示。开发了一种独特的支付机制, 满足众包

的两个重要设计原则。此外，拟议的机制进一步鼓励高素质的工人减少使用提示，这有助于确定和分配更多可能的报酬。在amazon机械土耳其人身上进行的实验表明，我们的方法确保了足够数量的低支出的高质量标签，并检测到高质量的工人。少

2018年3月6日提交;v1于2018年2月26日提交;最初宣布2018年2月。

## 102. 情感线: 多方对话的情感语料库

作者:陈胜国,许朝春, 郭传春,丁浩,黄, 顾伦伟

**摘要:** 感受情感是区分人和机器的一个重要特征。在所有用于情感检测的多模态资源中，文本数据集是那些除了语义之外包含的附加信息最少的数据集，因此被广泛用于测试所开发的系统。然而，大多数文本情感数据集只包含单个单词、句子或文档的情感标签，这使得讨论情感的语境流动具有挑战性。在本文中，我们介绍了情感线，第一个数据集的情感标签的所有话语中的所有发言只基于其文本内容。情感线对话是从朋友电视脚本和私人 facebook 信使对话收集。然后七种情绪中的一种，六个 ekman 的基本情绪加上中性的情绪，被 5 亚马逊 mturkers 标记在每个话语上。来自 2000 个对话的 29 245 个话语在 "情绪线" 中被标记为 "情感线"。本文还为情绪线上的情绪检测模型提供了几个强有力的基线。少

2018年5月30日提交;v1于2018年2月22日提交;最初宣布2018年2月。

### 103. 以网络为基础的 vr 实验

作者:小马,梅根·卡卡特,莱斯莉公园,埃里克·钱,莫尔·纳曼

文摘: 我们在虚拟现实 (vr) 设备和 web 技术日益普及的基础上, 利用众包技术在虚拟现实 (vr) 中进行行为实验。一种新的招聘和验证方法使我们能够创建一个由从亚马逊机械土耳其人招募的合格实验参与者组成的小组。使用这个面板, 我们进行了三个不同的众包虚拟现实实验, 每一个复制三个 vr 幻象之一: 地方错觉, 体现错觉, 和合理性错觉。我们对这些实验的经验和工人反馈表明, 使用众包进行基于网络的 vr 实验已经是可行的, 尽管仍然存在一些挑战——包括规模。这种众包虚拟现实实验在网络上有可能最终支持可复制的虚拟现实实验, 以低廉的成本与不同的人群。少

2018年2月26日提交;v1于2018年2月22日提交;最初宣布2018年2月。

### 104. 感知潜在因子模型: 评级预测与评价

作者:程志勇,应英,朱磊,莫汉·坎坎哈里

**摘要:** 虽然潜在因子模型（如矩阵分解）在评级预测方面具有良好的准确性，但它们会遇到几个问题，包括冷启动、非透明度和针对本地用户或项目的次优推荐。在本文中，我们使用带有分级的文本评审信息来解决这些限制。首先，我们在审查文本上应用了一个面向对象感知的主题模型（atm），从不同的角度对用户偏好和项目特征进行建模，并估计了用户对项目的方面重要性。然后将其重要性集成到一个新的方面感知潜在因素模型（alfm）中，该模型基于评级来学习用户和项目的潜在因素。特别是，alfm 引入了一个加权矩阵，将这些潜在因素与 atm 发现的同一组方面联系起来，以便将潜在因素用于估计方面评级。最后，通过方面评级的线性组合计算总体评级，这些评级由相应的方面重要性加权。为此，我们的模型可以缓解数据稀疏问题，获得良好的推荐可解释性。此外，一个方面的评级是加权的方面重要性，这取决于目标用户的喜好和目标项目的功能。因此，预计该方法可以在本地为每个用户项对更准确地建模用户对某个项目的首选项。对**来自亚马逊**和 yelp 2017 年挑战数据集的 19 个数据集进行了全面的实验研究。结果表明，与强基线方法相比，该方法取得了显著的改进，特别是对于收视率仅较少的用户。此外，我们的模型可以对推荐结果进行深入的解释。少

**2018 年 2 月 22 日提交;最初宣布 2018 年 2 月。**

## **105. 在机器学习中窃取超参数**

作者:王炳辉,龚振强

**摘要:** 超参数在机器学习中至关重要, 因为不同的超参数往往会导致性能明显不同的模型。超参数可能被视为机密, 因为它们具有商业价值, 并且学习者用来学习它们的专有算法是保密的。在这项工作中, 我们建议攻击窃取学习者学习的超参数。我们称之为攻击超参数窃取攻击。我们的攻击适用于各种流行的机器学习算法, 如脊回归、逻辑回归、支持向量机和神经网络。我们从理论和经验两方面评估攻击的有效性。例如, 我们评估我们对 **amazon** 机器学习的攻击。我们的结果表明, 我们的攻击可以准确地窃取超参数。我们还研究对策。我们的研究结果突出表明, 对于某些机器学习算法, 需要对我们的超参数窃取攻击进行新的防御。少

**2018 年 2 月 20 日提交;**v1 于 2018 年 2 月 14 日提交;**最初宣布** 2018 年 2 月。

## 106.snc: 一种使用实时编译进行符号数字计算的云服务平台

作者:张鹏,刘月明,邱美康

**摘要:** 云服务已广泛应用于 it 行业和科学研究。通过使用云服务, 用户可以将计算任务和数据从本地计算机移动到远程数据中心。通过通过轻量级和移动设备访问基于 internet 的服务, 用户可以在功能强大的计算机上部署多样化的云应用程序。科学计算领域这一范式的关键驱动因素包括强大的计算能力、按需配置和跨

平台互操作性。但是, 为了充分利用云服务进行科学计算, 我们需要设计一个特定于应用程序的平台, 以帮助用户高效迁移其应用程序。在此, 我们提出了一个用于符号-数值计算的云服务平台——snc。snc 允许云用户通过 c c++、python、java api 和 snc 脚本将任务描述为符号表达式。使用 llvmmn jvm 进行实时 (jit) 编译, 用于将用户代码编译为计算机代码。我们实现了 snc 设计, 并在几个流行的云平台 (包括线性最小化、蒙特卡罗积分、有限元组装和多体动力学) 上测试了广泛的符号数值计算应用 (包括非线性最小化、蒙特卡罗积分、有限元组装和多体动力学)。谷歌计算引擎、**亚马逊** ec2、微软 azure、rackspace、hp helion 和 vmware vcloud)。这些结果表明, 我们的方法可以跨多个云平台工作, 支持不同的语言, 并显著提高使用云平台进行符号数字计算的性能。这为激发在科学研究领域使用云计算进行符号数值计算的需要提供了一种途径。少

**2018 年 2 月 9 日提交;最初宣布** 2018 年 2 月。

## 107. 关于做还是作为? –关于让云应用程序对攻击者来说过于快速和愤怒的立场文件

**作者:**[nane kratzke](#)

**摘要:** 云应用程序还暴露了潜在或实际的漏洞——在服务端点旁边。而攻击者也有几个优势。他们可以选择武器、时间点和攻击点。

云应用程序安全工程工作通常集中在强化堡垒墙上，但很少假定攻击可能是成功的。因此，云应用程序依赖于它们的防御墙，但很少主动攻击入侵者。生物系统是不同的。他们承认防御 "墙" 可以在几个层面被破坏，因此利用主动和适应性防御系统来攻击潜在的入侵者——免疫系统。本立场文件提出了这样一种免疫系统启发的方法，以确保即使是未被发现的入侵者也能从云应用程序中清除。这使得入侵者在受害者系统上保持存在变得更加困难。使用流行的云服务基础架构（amazon web 服务、google 计算引擎、azure 和 openstack）进行的评估实验表明，这可以将未被检测到的入侵者的代理周期降至几分钟。少

2018 年 2 月 10 日提交;最初宣布 2018 年 2 月。

## 108. 实现轻量级多云 dsl，实现弹性和可转换的云原生应用程序

作者:[pterchristian quint](#), [nane kratzke](#)

**摘要:** 云原生应用程序是有意为云设计的，目的是利用云平台功能，如水平扩展和弹性——云平台带来的好处。除了经典的（通常是静态的）多层部署方案外，云原生应用程序通常还在更复杂但富有弹性的基础架构上运行。此外，还有一种趋势是使用弹性容器平台，如 kubernetes、docker 或 apache mesos。然而，尤其是多云用例是惊人的复杂处理。因此，云原生应用程序容易被供应商锁定。通常使用基于 tosa 的方法来解决这方面的问题



题。但是，这些应用程序拓扑定义方法在支持运行时云原生应用程序的多云适应方面是有限的。在本文中，我们分析了几种定义云原生应用程序在运行时可转移的方法。我们没有找到完全满足我们所有要求的方法。因此，我们引入了一个将弹性平台定义与云应用程序定义分离的解决方案。我们首先介绍了应用程序定义的特定于域的语言的注意事项，并在平台级别演示了评估结果，显示可以在不同的云服务提供商（如 azure）之间传输云原生应用程序和谷歌在几分钟内，而不是停机。评估涵盖了 amazon web 服务、microsoft azure、google 计算引擎和 openstack 提供的公共和私有云服务基础架构。少

2018 年 2 月 10 日提交;最初宣布 2018 年 2 月。

## 109. 利用编码技术加快分布式计算

作者:[konstantinos konstantinidis](#) , [aditya ramamoorthy](#)

**抽象:** 利用分布式计算框架（如 mapreduce）的大型群集通常会处理按 pb 或更多的顺序运行的数据。数据的绝对大小使数据无法在一台计算机上进行处理。这些方法的理念是将整个作业划分为在不同服务器上执行的较小任务;这就是所谓的地图阶段。接下来是数据洗牌阶段，在服务器之间交换适当的数据。最后所谓的减少阶段，完成计算。在以前的工作中探索的减少整体执行时间的一个潜在方法是在计算和通信之间进行自然权衡。具体来说，

我们的想法是运行放置在明智选择的服务器上的地图任务的冗余副本。洗牌阶段利用节点的位置并利用编码传输。这种方法的主要缺点是，它要求将原始作业拆分为在系统参数中呈指数级增长的多个映射任务。这是有问题的，因为我们证明，将作业拆分得太细实际上会对整个执行时间产生不利影响。在这项工作中，我们表明，一个人可以同时获得较低的通信负载，同时确保工作不需要拆分得太细。我们的方法揭示了这个问题和一类称为可解析设计的组合结构之间的深层关系。对可解析设计的适当解释可以在分裂水平比以前的工作低指数的情况下开发编码的分布式计算方案。我们为一种众所周知的分布式算法（terasort）提供了在 **amazon ec2** 聚类上获得的实验结果。我们获得超过 4.69X 在基准方法的加速和超过 2.6X 在目前的最先进的状态。少

**2018 年 2 月 8 日提交;最初宣布** 2018 年 2 月。

## 110. gosig: 关于区块链的对抗性广域网的可扩展拜占庭共识

**作者:**李培伦,王国赛,陈晓琪, 徐伟

**摘要:** 现有的拜占庭容错 (bft) 协议在联盟区块链方案中面临重大挑战。一方面，我们可以对底层互联网的可靠性和安全性做很少的假设。另一方面，联盟区块链上的应用程序需要像 bit-coin 这样的可扩展系统，但提供更高的性能以及可证明的安全性。我们提出了一个新的 bft 协议 gosig，它将协议层中基于密码的秘

密引线选择和多轮投票与实现层优化（如基于基于文件的消息传播）结合起来。特别是, gosig 保证了即使在完全由对手控制的网络中的安全性, 同时通过易于实现的网络连接假设提供可证明的活力。在由横跨五大洲 14 个城市的 140 台 **amazon** ec2 服务器组成的广域测试平台上, 我们展示了 gosig 每秒可以在不到 1 分钟的交易确认时间内实现 4, 000 多笔交易。少

2018 年 2 月 5 日提交;最初宣布 2018 年 2 月。

## 111. 360 视频时空高光检测的深度排名模型

作者:y 细 [yaeyu](#), [sangholee](#), [joonil na](#), [jaeyunkang](#), [gunhee kim](#)

**摘要:** 我们通过从空间和时间上对 360 度视频进行总结来解决 360 度视频中的高光检测问题。鉴于一个长的 360 度视频, 我们从 360 度视频的无限视野 (fov) 空间上选择看起来愉快的正常视场 (nfov) 段, 并将其临时总结为一个简洁和信息丰富的亮点, 作为一个选定的子集子镜头。提出了一种新的深部排名模型——合成视图评分 (cvs) 模型, 该模型生成每个视频段的合成球面评分图, 并确定哪种视图在推理时适合通过滑动窗口内核突出显示。为了评估建议的框架, 我们对 Pano2Vid 基准数据集以及我们从 youtube 和 vimeo 新收集的 360 度视频高亮数据集进行了实验。通过使用定量摘要指标和通过 **amazon** 机械土耳其人进行的用户研究进行评估, 我们证明我们的方法优于几种最先进的突出检测

方法。我们还表明, 我们的模型在推理上比 autcam 快 16 倍, autcam 是最早的 360 度视频的汇总算法之一。

2018 年 1 月 31 日提交;最初宣布 2018 年 1 月。

## 112. 建议的多指针协同注意网络

作者: [易泰](#), [卢安团](#), [小张辉](#)

**摘要:** 许多最近最先进的推荐系统, 如 d-att、trannet 和 deepcon, 利用评审来进行表示学习。本文提出了一种新的神经结构, 用于推荐和评价。我们的模型基于多层次的范式, 并基于这样的直觉, 即并非所有的评论都是平等的, 即只有少数人是重要的。但是, 应根据目前的目标动态推断重要性。为此, 我们提出了一个基于表的审查学习方案, 该方案提取重要的评论, 然后以逐字逐句的方式对其进行匹配。这不仅可以将信息最丰富的评论用于预测, 还可以更深入地进行文字层面的交互。我们基于指针的方法使用一种新的基于 gumbel-softmax 的指针机制, 该机制能够将离散向量集成到可微的神经架构中。我们的指针机制本质上是共同关注的, 学习指针是共同依赖于用户项关系的。最后, 我们提出了一个多指针学习方案, 学习结合用户和项目之间交互的多个视图。总体而言, 我们通过对来自 amazon 和 yelp 的 \textbf{24} 基准数据集进行的大量实验来展示我们提出的模型的有效性。实验结果表明, 与 trannet 和 deepcon 相比, 我们的

方法明显优于现有的最先进技术，相对改善率分别高达 19% 和 71%。我们研究了我们的多指针学习机制的行为，揭示了基于评审的推荐系统中的证据聚集模式。少

**2018 年 6 月 21 日提交;**v1 于 2018 年 1 月 28 日提交;**最初宣布** 2018 年 1 月。

### 113. 了解隐私态度与解锁身份验证之间的联系

**作者:**adam j. aviv, ravi kuber

**抽象:** 在本研究中，我们将研究用户对与移动设备相关的隐私和安全的态度以及存储在移动设备上的数据可能会影响解锁身份验证的强度的方式，重点是 android 的图形解锁模式。我们与**亚马逊机械土耳其人**进行了在线研究 (n=750) 使用自报解锁身份验证选项，以及 likert 规模协议/对一组七个隐私/安全提示的分歧响应。然后，我们分析了多个维度的响应，包括 likert 响应的直线平均值，以及使用主成分分析来揭示潜在因素。我们发现，对七个问题中的两个问题的答复证明是相关和重要的。这两个问题考虑了对移动设备上存储的数据的普遍关切的态度，以及对已知行为者未经授权访问的关切的态度。不幸的是，对于揭露解锁身份验证强度之间的联系的有效性的更广泛的一组问题的有效性，无法得出更大的结论 (pearson rankr=--0.08, p<0.1)。然而，我们的两个因素解决方案暴露了在对人口群体的反应差异，包括年龄，性别

和居住类型。这项研究的结果表明,对移动设备隐私安全的看法与其中的感知威胁之间可能存在联系,但还需要进行更多的研究,特别是在开发更好的调查和测量技术方面。与移动设备具体相关的隐私安全态度。少

2018 年 2 月 1 日提交;[v1](#) 于 2018 年 1 月 23 日提交;[最初宣布](#) 2018 年 1 月。

#### 114. 价值存储中的乐观执行

作 者 :[duong nguyen](#), [aleksey charapko](#), [sandeep kulkarni](#), [murat demirbas](#)

**摘要:** cap 定理的局限性意味着,如果在存在网络分区的情况下需要可用性,则必须牺牲顺序一致性,这是一种更自然的系统设计一致性模型。我们关注的问题是,如果设计人员的算法具有顺序一致性,但却面临一个底层键值存储,该存储提供了较弱(例如,最终或因果)的一致性,则她应该做什么。我们提出了一种基于检测回滚的方法:设计人员标识一个正确性谓词(如  $p$ ),并继续运行协议,因为我们的系统监视  $p$ 。如果违反  $p$ (因为基础键值存储提供了较弱的一致性),系统将回滚并在  $p$  保持的状态下恢复计算。我们在伏地魔键值存储中评估了这种方法。我们在 [amazon aws](#) 上部署伏地魔的实验表明,与顺序一致性相比,使用最终与监控的一致性可以提高 20–40% 的吞吐量。我们还表明,

监视器本身的开销很小（通常小于 8%），检测冲突的延迟非常低。例如，在不到 1 秒的时间内检测到 99.9% 以上的违规情况。少

2018 年 1 月 22 日提交;最初宣布 2018 年 1 月。

## 115. 深度强化学习聊天机器人 (短版)

作者 :[iulian v. serban](#), [chinnadhuraisankar](#), [mathieu germain](#), [sazhengzhang](#) , [zh 心 han lin](#), [sandeepsubramanian kim](#), [michael pi 片断 er](#) , [sarath chandar](#) , [nan rosemary ke](#), [sai rajeswar](#), [亚历山大 de brebisson](#), [jose m. r.sotelo](#), [dendi suhubdy](#), [vincent michalski](#), [亚历山大 nguyen](#), [joelle pineau](#) , [yeshua bengio](#)

**摘要:** 我们介绍 milabot: 一个深入强化学习聊天机器人开发的蒙特利尔学习算法研究所 (mila) 为**亚马逊亚历克莎**奖比赛。milabot 能够通过语音和文本与人类就流行的小谈话话题进行交谈。该系统由一系列自然语言生成和检索模型组成，包括神经网络和基于模板的模型。通过将强化学习应用于众包数据和现实世界中的用户交互，该系统接受了培训，以便从其集成中的模型中选择适当的响应。通过对实际用户进行的 aeb 测试对该系统进行了评估，该系统的性能明显好于其他系统。研究结果突出了将集成系统与深度强化学习结合起来的潜力，作为开发现实世界开放域会话代理的一条富有成效的途径。少



2018 年 1 月 20 日提交;最初宣布 2018 年 1 月。

## 116. 分布式图形分析的编码计算

作者 :[saurav prakash](#), [amirhoshussein reisizadeh](#), [ramtin Pedarsani](#) , [salman avestimehr](#)

**摘要:** 为了应对对大规模图形结构数据集高效处理日益增长的需求,最近开发了许多分布式图形计算系统。由于这些系统在计算的每一步都需要在计算机之间交换许多消息,因此通信带宽被观察到是一个主要的性能瓶颈。我们提出了一个编码计算框架,它系统地注入计算阶段的冗余,以便在通信阶段实现编码机会;从而大大减少通信负载。具体而言,我们提出了一种编码方案,该方案允许 erdős-rényi 图在计算负载和平均通信负载之间进行(渐近)反向线性权衡。随着图形大小的增加,该方案被证明是最优的渐近方案。对于有限尺寸的图形,我们通过数值分析证明,对于给定的计算负载  $R$ , (即当每个图形顶点都被仔细地存储在  $R$  服务器),该方案将平均通信负载削减了一个乘法因子。此外,我们将我们的结果推广到其他三个随机图模型——随机二部模型、随机块模型和功率定律模型。特别是,我们证明了我们的方案在这些流行的图形模型中,也实现了分布图处理中计算负载和通信负载之间的反向线性权衡。我们终于在 **amazon ec2** 集群上进行了实验,以实际展示我们的编码方案的影响。特别是,对于流行的年龄排



名算法, 我们证明了我们提出的编码方案的速度高达 47%与基准方法的比较。少

2018 年 7 月 4 日提交;v1 于 2018 年 1 月 16 日提交;最初宣布 2018 年 1 月。

## 117. 会话代理的评价与比较

**作者:** [anu venkatesh](#), [chandra khatri](#), [ashwinram](#), [fef-guo](#), [raefer gabriel](#), [ashish nagar](#), [rohit prasad](#), [ming cheng](#), [behnam hedayatnia](#), [angelikimetallinou](#), [rahul goel](#), [shaohua yang](#), [anirudh raju](#)

**摘要:** 对话代理的受欢迎程度正在不断提高。然而, 尽管近年来研究兴趣显著增加, 但在非目标导向对话领域仍有许多工作要做。为了在对话 ai 中的最新水平, 亚马逊发起了亚历克莎奖, 这是一项价值 250 万美元的大学竞赛, 16 个被选中的大学团队建立了对话代理, 以提供最好的社交对话经验。亚历克莎奖为学术界提供了一个独特的机会, 可以使用数百万用户使用的现场系统进行研究。与评价对话相关的主观性是建立非目标对话系统这一挑战的关键因素。本文提出了一种综合评价策略, 采用多种指标, 通过选择与人的判断相关联的指标来降低主观性。建议的度量提供了对会话代理的细粒度分析, 而这在人工评分中没有捕获。我们表明, 这些指标可以作为人类判断的合理代名词。我们提供了一个机制来统一选择表现最佳的代理商的指标, 这也已应用于整个亚历克莎奖的比赛。据我们所知, 迄今为止, 它是评估具有数百万次对话

和几十万用户评分的代理的最大设置。我们相信这项工作是迈向对话认可机构自动评估过程的一个步骤。少

2018 年 1 月 10 日提交;最初宣布 2018 年 1 月。

## 118. 对话 ai: 亚历克莎奖背后的科学

**作者:**ashwin ram, rohit prasad, chandra khatri, anu venkatesh ,  
raefergabriel, chingliu, jeff nunn , behnam  
hedayatnia, mingcheng, ashish nagar , eric king, katebland,  
amanda wartick, yi pan, han song, sk Jayadevan, gene hwang,  
art pettigruue

**摘要:** 对话代理的受欢迎程度正在不断提高。然而, 在社会对话以及在广泛的领域和主题上的自由形式对话领域, 仍有许多工作要做。为了在对话 ai 中的最新水平, **亚马逊**发起了亚历克莎奖, 这是一项价值 250 万美元的大学竞赛, 16 个选定的大学团队在比赛中受到挑战, 要求他们建立对话代理, 即所谓的社交机器人。在体育、政治、娱乐、时尚和技术等热门话题上与人类进行连贯和引人入胜的交谈, 持续 20 分钟。亚历克莎奖为学术界提供了一个独特的机会, 可以使用数百万用户使用的实时系统进行研究。比赛为大学团队提供了大规模的真实用户对话数据, 以及用户提供的评级和反馈, 并由亚历克莎团队进行了注释。这使团队能够在整个竞争过程中有效地迭代和进行改进, 同时通过实时用户交互进行实时评估。为了建立他们的社会机器人, 大学团队结合了最先进的技术和新的策略领域的自然语言理解, 上下文建模, 对话管

理, 响应生成, 和知识获取。为了支持参与团队的努力, 亚历克莎奖团队进行了大量的科学和工程投资, 以建立和改进对话语音识别、主题跟踪、对话评估、语音用户体验和流量工具管理和可扩展性。本文概述了大学团队以及亚历克莎奖团队为实现解决会话 ai 问题的共同目标所创造的进步。少

**2018 年 1 月 10 日提交;最初宣布** 2018 年 1 月。

## 119. 与人群一起探索陈规定型观念和偏见数据

**作者:**胡泽源,朱莉娅·斯特劳特

**摘要:** 我们研究的目的是提供信息, 说明人群在预测可能在研究人员不知情的情况下对数据集产生偏见的陈规定型观念有多大的用处。人群预测的结果有可能在数据收集过程中使用, 以帮助防止可疑的陈规定型观念对数据集引入偏差。我们通过**要求亚马逊机械土耳其人 (amt)** 上的人群完成两项类似的人类智力任务 (hit) 来进行研究, 提出与他们的个人经历相关的陈规定型观念。我们对这些答复的分析侧重于确定工人建议的多样性程度及其人口结构。通过这个过程, 我们开始讨论人群在机器学习数据收集解决这一难题的有用程度。少

**2018 年 1 月 10 日提交;最初宣布** 2018 年 1 月。

## 120. 电子商务的生成性对抗网络

作者:asutosh kumar, arijit bis 遭到,subhajit sanyal

**摘要:** 亚马逊、阿里巴巴和 flipkart 等电子商务公司每年处理数十亿美元的订单。然而, 这些订单只占有所有可信订单的一小部分。探索所有合理订单的空间可以帮助我们更好地了解电子商务生态系统中各个实体之间的关系, 即客户及其购买的产品。在本文中, 我们提出了一个生成对抗性网络 (gan) 在电子商务网站上的订单。一旦训练好, gan 中的生成器就可以生成任意数量的合理订单。我们的贡献包括: (a) 对电子商务订单进行密集和低维的表示; (b) 培训一个有真实订单的电子商务 (ecgan), 以显示拟议范式的可行性; (c) 培训一个电子商务条件的 gan ( $ec^2GAN$ ) 生成涉及特定产品的合理订单。我们提出了几种定性的方法来评价 ecgan, 并证明了它的有效性。该  $ec^2GAN$  用于对刚刚引入电子商务系统的产品的可能订单进行各种定性。在大多数情况下, 拟议的方法  $ec^2GAN$  明显优于基线。少

2018 年 1 月 10 日提交;最初宣布 2018 年 1 月。

## 121. 基于异构侧向信息融合的推荐系统学习

作者:赵欢,姚全明,宋阳秋,郭德伦

**文摘:** 基于协同过滤 (cf) 的方法已成为推荐系统 (rs) 中最常用的技术。近年来, 在 cf 中引入了用户之间的社会关系和项目元数据等各类副业信息, 并证明对提高推荐业绩有效。此外, 侧面信

息可以缓解传统 cf 面临的数据稀疏和冷启动问题, 但以前的作品分别处理不同类型的信息, 从而丢失了不同类型的侧可能存在的信息信息。在本文中, 我们研究了异构信息网络 (hin) 的应用, 它提供了不同类型信息的灵活表示, 以增强基于 cf 的推荐方法。由于 hin 可以是一个复杂的图形, 表示两种实体类型之间存在的多种类型的关系, 因此我们需要解决基于 hin 的 rs 所面临的两个具有挑战性的问题: 如何捕获驱动用户和中的项之间相似性的复杂语义。hin, 以及如何融合异构信息来支持推荐。为了解决这些问题, 我们建议将半程图应用于基于 hn 的 rs, 并利用 "矩阵分解 (mf) + 分解机 (fm)" 框架解决信息融合问题。对于 mf 部分, 我们从每个半图中获取用户项相似矩阵, 然后应用低阶矩阵逼近来获得用户和项目的潜在特征。对于 fm 部分, 我们建议在从 mf 部分获得的特征上应用带有组套索 (fmg) 的 fm, 以训练推荐模型, 同时确定有用的元图。在两个大型实际数据集 (amazon 和 yelp) 上的实验结果表明, 我们提出的方法优于 fm 和其他最先进的基于 hn 的推荐方法。少

**2018 年 10 月 9 日提交;**v1 于 2018 年 1 月 8 日提交;**最初宣布** 2018 年 1 月。

## 122.slugbot: 一种新颖且可扩展的开放领域社会机器人框架的应用

**作者:**kevin k. bowden, jiqiwu, shereen oraby, amita misra, marilyn walker

**摘要:** 在本文中, 我们介绍了一个新颖的, 开放领域的社交机器人亚马逊亚历克莎奖比赛, 旨在进行友好的对话与用户的各种主题。我们展示了我们的模块化系统, 突出了我们不同的数据源, 以及我们如何将人的思维作为数据管理的模型。此外, 我们还建立和使用自然语言理解和信息检索工具及 api, 以扩大我们的知识基础。我们描述了我们的半结构化、可扩展的框架, 用于构建特定主题的对话流, 并详细介绍了我们的对话管理计划和评分机制。最后, 我们简要评估了我们系统的性能, 并观察了开放领域 socialbot 所面临的挑战。少

**2018 年 1 月 4 日提交;最初宣布** 2018 年 1 月。

### 123. 在 2d 蜜蜂蜂巢中实现密集物体跟踪

**作者:** [ketarzyna bozek](#), [laetitia hebert](#) , [亚历山大 s mikheyev](#), [greg j stephens](#)

**摘要:** 从人类人群到组织中的细胞, 检测和有效跟踪密集构型中的多个物体是一个重要的、尚未解决的问题。过去, 图像分析的局限性限制了对密集群体的研究, 限制了对单个或子集标记个体的跟踪, 或限制了粗粒度的群体水平动态, 所有这些都产生了不完整的信息。在这里, 我们将卷积神经网络 (cnn) 与蜜蜂蜂巢的模型环境结合起来, 从原始图像数据中自动识别密集组中的所有个体。我们创建新的, 适应的个人标签, 并使用分割架构 u-net 的损失函数取决于对象的身份和方向。此外, 我们还以反复的方

式利用视频录制的时间规律性, 实现了近乎人类水平的性能, 同时与原始 u-net 架构相比, 将网络大小减少了 94%。鉴于我们对 cnn 的新应用, 我们生成了广泛的特定于问题的图像数据, 其中标记的示例是通过与 **amazon** 机械土耳其的自定义接口生成的。此数据集包含超过 37, 000 个标记蜜蜂实例, 在 720 个视频帧中位于 2 个 fps, 这是开发和测试跟踪方法的广泛资源。我们正确地检测到 96% 的人的位置误差约 7% 的典型身体尺寸, 方向误差为 12 度, 接近人类评分者的变异性。我们的研究结果允许在一个网络体系结构中有效地准确地确定跨时间序列图像数据的对象位置和方向, 从而为实现基于图像的密集对象跟踪迈出了重要一步。少

2017 年 12 月 22 日提交;最初宣布 2017 年 12 月。

## 124. 具有社会对话排名的合奏模型

作者 :[ioannis papaioannou](#), [amanda cercas curry](#), [jose l.part](#), [igor shalyminov](#), [exinnuo xu](#), [yanchao yu](#), [Ondř ej duš ek](#), [verena rieser](#), [oliver leemon](#)

**摘要:** 开放领域社会对话是人工智能的长期目标之一。今年,亚马逊亚历克莎奖的挑战首次宣布, 真正的客户获得了全球领先大学开发的系统评级。挑战的目的是 "连贯和引人入胜地与人在热门题目交谈 20 分钟"。我们描述了我们的亚历克莎奖系统 (称为 "alana"), 该系统由一群机器人组成, 结合了基于规则的系统 and 机

器学习系统, 并使用上下文排名机制来选择系统响应。护林员接受了比赛期间收到的真实用户反馈的培训, 解决了如何在比赛中获得的嘈杂和稀疏反馈进行培训的问题。少

**2017 年 12 月 20 日提交;最初宣布** 2017 年 12 月。

## 125. 亚马逊机械土耳其人收入的数据驱动分析

**作者:**[kotaro hara](#), [abi adams](#), [k 饿 miland](#), [saiph savage](#), [chris callison-burch](#), [jeffrey bigham](#)

**摘要:** 越来越多的人作为在线人群工作的一部分工作, 其特点是工资低;然而, 我们对工资分配和低收入的原因了解甚少。我们记录了 2, 76 名工人在**亚马逊机械土耳其人**身上执行 380 万项任务。我们的任务水平分析显示, 工人每小时的工资中值只有 ~ 2\$, 只有 4% 的工资超过 7.25 美元. \$平均请求者支付的费用超过\$, 尽管报酬较低的请求者承担的工作要多得多。我们的工资计算受到无报酬工作如何包含在我们的工资计算中的影响, 例如, 花在寻找任务、处理被拒绝的任务上的时间以及从事最终未提交的任务的时间。我们进一步探讨了每小时工资较高的任务和工作模式的特点。我们的分析为未来的平台设计和工人工具提供了信息, 为人群工作创造了更积极的未来。少

**2017 年 12 月 28 日提交;v1** 于 2017 年 12 月 14 日提交;**最初宣布** 2017 年 12 月。



## 126. 协同绘图: 协作绘图的可视化对话框

作者 : 金 华      kim,    devi    parikh    ,    dhruv  
batra, byoung-takzhang, yuend 对待    tian

**摘要:** 在这项工作中, 我们提出了一个目标驱动的协作任务, 其中包含虚拟环境中的愿景、语言和行动作为其核心组件。具体来说, 我们开发了两个代理之间的协作 "图像绘制" 游戏, 称为 codraw。我们的游戏是建立在一个虚拟世界, 其中包含可移动的剪贴画对象。两名球员, 泰勒和抽屉, 参与其中。泰勒在语义上有意义的配置中看到包含多个剪贴画的抽象场景, 而抽屉则尝试使用可用的剪贴画在空画布上重建场景。这两个玩家使用自然语言通过双向交流进行交流。我们收集 ~ 10k 对话框的电子绘制数据集, 其中包括一个柜员和一个抽屉之间交换的 138k 消息, 从亚马逊机械土耳其人 (amt)。我们分析我们的数据集, 并提出三个模型来模拟玩家的行为, 包括一个注意模型, 以描述和绘制多个剪贴画在每一轮。将注意模型与其他模型进行了定量比较, 以显示传统方法如何适用于这一新任务。我们还展示了定性可视化。少

2017 年 12 月 15 日提交;最初宣布 2017 年 12 月。

## 127. 从嘈杂的单一标记数据中学习

作者: ashish khetan, zachary c. lipton, anima anandkumar

**摘要:** 监督学习依赖于注释的示例, 这些示例被视为 "接地真理"。但这些标签往往来自嘈杂的众包平台, 如**亚马逊**机械土耳其人。从业者通常收集每个示例的多个标签, 并汇总结果以减轻噪音 (典型的众包问题)。给定固定的注释预算和无限限制的未标记数据, 冗余注释以牺牲较少的标记示例为代价。这就提出了两个基本问题: (1) 我们如何才能最好地向嘈杂的工人学习? (2) 我们应该如何分配我们的标签预算, 以最大限度地提高分类器的性能? 提出了一种新的算法, 用于从嘈杂的众包数据中对标签和工人质量进行联合建模。交替最小化依次进行, 从与当前模型的不一致中估计工人质量, 然后通过优化占当前工人质量估计的损失函数来更新模型。与以前的方法不同, 即使每个示例只有一个注释, 我们的算法也可以估计工作人员的质量。我们为使用我们的算法学习的模型建立了一个泛化误差, 并从理论上确定, 当工人质量超过阈值时, 最好给许多例子贴上一次标签 (相对于乘法较少)。在 imagenet (模拟嘈杂的工作人员) 和 ms-coco (使用真正的众包标签) 上进行的实验证实了我们算法的优点。少

**2018年5月20日提交;**v1 于 2017 年 12 月 12 日提交;**最初宣布** 2017 年 12 月。

## 128. 家庭数字语音助理的不安全-亚马逊亚历克莎作为一个案例研究

**作者:**雷新宇,杜冠华, 刘永华,李志宇,田谢

**摘要:** 近年来, 家庭数字语音助理 (hdva) 越来越受欢迎。用户可以使用语音控制智能设备, 并通过这些 hdva (例如, **amazon** 亚历克莎、google home) 获得生活帮助。在这项工作中, 我们研究了 hdva 服务的不安全使用亚马逊亚历克莎作为一个案例研究。我们披露了三个安全漏洞, 这些漏洞的根源在于亚历克莎服务的不安全访问控制。然后, 我们利用他们设计两个概念证明攻击, 家庭入室盗窃和假订单, 在那里, 对手可以远程命令受害者的亚历克莎设备打开一扇门或从 amazon. com 下订单。不安全的访问控制是, 亚历克莎设备不仅依赖于单因素身份验证, 而且还接受语音命令, 即使周围没有人。因此, 我们认为 hdva 应该有另一个身份验证因素, 即基于物理存在的访问控制;也就是说, 只有当附近检测到任何人时, 他们才能接受语音命令。为此, 我们设计了一个虚拟安全按钮 (vsbutton), 它利用 wifi 技术来检测室内人体的运动。一旦检测到任何室内人体运动, 亚历克莎设备就可以接受语音命令。我们的评价结果表明, 它可以有效地区分室内运动和无运动和室外运动的情况下, 在实验室和现实世界的设置。

少

**2018 年 4 月 8 日提交;**v1 于 2017 年 12 月 8 日提交;**最初宣布** 2017 年 12 月。

## 129. 项目推荐和生成的兼容性家庭学习

**作者:**石永祥, 张开岳, 孙天林, 孙敏

**摘要:** 物品（如衣服和鞋子）之间的兼容性是客户采购决策中的一个主要因素。然而，学习 "兼容性" 具有挑战性，因为 (1) 比相似性更广泛的兼容性概念, (2) 兼容性的不对称性质, (3) 只观察到一小部分兼容和不兼容的项目。我们提出了一个端到端可训练系统，将每个项目嵌入到一个潜在的向量中，并将查询项投影到同一空间中的  $k$  兼容原型中。这些原型反映了兼容性的广泛概念。我们将嵌入和原型称为 "兼容性系列"。在我们学到的空间中，我们引入了一个新的投影兼容性距离 (pcd) 函数，该函数是可微的，并通过目标实现至少一个原型接近兼容项目来确保多样性，而没有一个原型接近不兼容的项目项目。我们在玩具数据集、两个亚马逊产品数据集和 polyvore 装备数据集上评估我们的系统。我们的方法始终实现最先进的性能。最后，我们展示了我们可以可视化候选兼容原型使用元数据正则化的条件生成抗性网络 (mrcgan)，其中输入是投影原型，输出是兼容项的生成图像。我们要求人类评估人员判断我们生成的图像与直接基于查询项的 cgans 生成的图像之间的相对兼容性。我们生成的图像是非常可取的，选票数量大约其他人的两倍。少

**2017 年 12 月 1 日提交;最初宣布** 2017 年 12 月。

### 130. 领域适应的图像转换

**作者 :** zak murez, soheil Kolouri, david kriegman, ravi ramamoorthi, kyungnam kim

**摘要:** 我们提出了一个无监督域适应的通用框架, 该框架允许在源域上训练的神经网络在不同的目标域上进行测试, 而无需在目标域中进行任何训练注释。这是通过添加额外的网络和损耗来实现的, 这些网络和损失有助于规范主干编码器网络提取的功能。为此, 我们提出了一种新的使用新的无配对图像到图像转换框架来约束编码器网络提取的特征。具体来说, 我们要求提取的要素能够重建两个域中的图像。此外, 我们还要求从两个域中提取的图像中提取的要素的分布是无法区分的。最近的许多工作可以看作是我们总体框架的具体案例。我们在分类任务中应用 mnist、usps 和 svhn 数据集与 amazon、webcam 和 dslr office 数据集之间的域适应方法, 以及 gta5 和 cityscapes 数据集之间的域适应, 以执行细分任务。我们展示了每个数据集的最先进的性能。少

**2017 年 12 月 1 日提交;最初宣布** 2017 年 12 月。

### 131. 双面设施位置

**作者 :** [reza alijani](#), [Siddhartha banerjee](#) , [sreenivas gollapudi](#), [kostas kollias](#), [kamesh munagala](#)

**摘要:** 近年来, 许多成功的电子商务市场平台兴起, 如亚马逊市场、bnb 航空公司、uber/lyft 和 up 工作, 其中一个中央平台可调解买卖双方之间的经济交易。在这些平台的推动下, 我们制定了一套设施选址问题, 我们将双方设施的位置称为双面设施的位

置。在我们的模型中，代理到达基础度量空间中的节点，其中任何买家和卖家之间的度量距离捕获相应匹配的质量。该平台在节点上发布价格和工资，并打开一套设施，将代理商路由到。任何设施的代理都被假定为匹配。该平台通过在节点和路由到的设施之间施加距离约束来确保高质量的匹配质量。它通过确保流向设施的流量至少是预先指定的下限来确保高服务可用性。在这些制约因素的情况下，该平台的目标是在预算余额疲软的情况下，即利润为非负，最大限度地增加社会盈余（或贸易收益）。我们提出了一个近似算法，这个问题，产生一个 $(1+\epsilon)$ 任何常数的盈余近似值  $\epsilon > 0$ ，同时以恒定的因子放松比赛质量（即任何匹配的最大距离）。我们使用一个  $lp$  舍入框架，很容易扩展到其他目标，如最大限度地增加交易量或利润。我们通过考虑动态的市场环境来证明我们的模型是合理的，在这种环境中，代理根据随机过程到达，并且对匹配具有有限的耐心（或截止日期）。我们执行排队分析，以表明对于将代理路由到设施并为之匹配的策略，确保代理的低放弃概率减少到确保每个设施都有足够的流量到达。少

**2018年2月27日提交;**v1于2017年11月30日提交;**最初宣布** 2017年11月。

## 132. 科学驱动的创新为移动产品提供动力：云 ai 与智能设备上的设备 ai 解决方案

作者:[德广港](#)

**摘要:** 近年来, 移动设备 (如 iphone) 因其给人类生活带来的便利而日益普及。一方面, 来自异构信息源的丰富的用户分析和行为数据 (包括每个应用级别、应用交互级别和系统交互级别) 使提供更好的服务 (如推荐、广告) 成为可能) 面向客户, 这进一步推动了了解用户行为和提高用户参与度的收入。为了取悦客户, 智能个人助理 (如**亚马逊**亚历克莎, 谷歌家居和谷歌 now) 是非常可取的, 以提供实时音频, 视频和图像识别, 自然语言理解, 舒适的用户互动接口, 满意的推荐和有效的广告定位。本文介绍了我们在移动设备上开展的研究工作, 其目的是通过利用统计和大数据科学、机器学习和深度学习、用户建模和营销技术, 提供更智能、更方便的服务。在移动设备上带来显著的用户增长和用户参与度 (和满意度)。开发的新功能建立在云端或设备侧, 和谐地协同工作, 以提高当前的服务, 目的是提高用户的满意度。我们用不同的案例研究从系统和算法的角度来设计这些新功能, 通过这些案例研究, 人们可以很容易地了解科学驱动的创新如何有助于提供更好的技术服务, 带来更多的收入在业务中的发展。同时, 这些研究工作也有明显的科学贡献, 并在顶级场馆发表, 这些场馆对移动 ai 产品发挥着越来越重要的作用。少

**2017 年 11 月 20 日提交;最初宣布** 2017 年 11 月。

### **133.隐私是怎么回事? : 智能个人助理中的用户偏好和隐私问题**

**作者:** [lydia manikonda](#), [aditya deotale](#) , [subbarao kambampati](#)



**摘要:** 人工智能 (ai) 最近的突破使个人能够出于各种原因依赖自动化系统。其中一些系统是目前流行的支持语音的系统, 如**亚马逊的** echo 和 home 的谷歌, 也被称为智能个人助理 (ipa)。尽管人们对隐私和道德影响表示关切, 但这些投资促进机构的用户似乎仍在继续使用这些系统。我们的目标是调查用户在使用 ipa 时关注隐私以及如何处理这些问题。通过利用网上发布的评论以及对调查的答复, 本文提供了一套与用户兴趣和隐私挑战有关的检测标记的见解。这些见解表明, 这些系统的用户, 无论其对隐私的关切如何, 在日常生活中利用投资促进机构方面总体上都是积极的。然而, 有相当大比例的用户关心隐私, 并采取进一步行动解决相关问题。有一定比例的用户表示他们没有任何隐私问题, 但当他们了解到这些设备的 "总是听" 功能时, 他们对隐私的担忧就增加了。少

**2017 年 11 月 20 日提交;最初宣布** 2017 年 11 月。

#### 134. 随机游走重组系统的调查

**作者:**[拉克拉纳特学期](#)

**摘要:** 推荐引擎已成为当今电子商务系统中不可或缺的组成部分。从在**亚马逊**推荐书籍到在脸谱等社交网络中寻找朋友, 它们已经无处不在。通常, 推荐系统可分为两大类: 基于内容的模型和基于协同过滤的模型。这两个模型都在用户和项之间建立关系, 以提



供建议。基于内容的系统通过利用从可用上下文中提取的要素来实现此任务，而协作系统则使用用户项子集之间的共享兴趣。还有另一种相对未探索的方法来提供建议，该方法利用了一个名为随机游走的随机过程。本研究是一项调查，探讨了推荐系统中随机游走的使用案例，并试图对其进行分类。少

2017 年 11 月 11 日提交;最初宣布 2017 年 11 月。

### 135. 对人体动作射击吸引力的认识与预测

作者:戴斌,王宝源,姜华

**摘要:** 从人类动作拍摄序列中选择有吸引力的照片是相当具有挑战性的，因为“吸引力”的主观性质，这主要是人类动作姿态和背景的综合因素。以前的作品一直在积极研究高层次的图像属性，包括有趣，记忆，大众化，和美学。然而，他们中没有一个人研究过人类动作镜头的“吸引力”。在本文中，我们提出了第一次研究人类动作镜头的“吸引力”，采取了系统的数据驱动的方法。具体来说，我们创建了一个由大约 8000 张高质量动作拍摄照片组成的新的动作拍摄数据集。我们进一步对亚马逊机械土耳其人 (amt) 进行了丰富的众包人类判断研究，研究了一张照片的全球吸引力，以及一张照片的相对吸引力。进一步提出了一种具有新的混合分布匹配损耗的深暹罗网络，以充分利用这两种类型的评级。广泛的实验表明: (1) 动作镜头吸引力的性质是主观的，但可以预测的;

2) 我们提出的方法对预测有吸引力的人类动作镜头既有效又有效。少

2017 年 11 月 2 日提交;最初宣布 2017 年 11 月。

### 136. 只需问: 构建可扩展自助口语理解的体系结构

作者 :anjishnu kumar, arpit gupta, julian chan, sam tucker, bjorn hoffmeister, markus dreier, stanislav peshterliev, ankur ganhe, denis filiminov, ariya rastrow, christian monson , agnika kumar

**摘要:** 本文介绍了作为亚历克莎技能工具包 (ask) 基础的机器学习体系结构的设计, 该工具集是一种大规模的口语理解 (slu) 软件开发工具包 (sdk), 使开发人员能够扩展 **amazon** 的功能。虚拟助理亚历克莎在**亚马逊**, 基础设施为通过 ask 以及 **aws** 的 **amazon** lex slu 服务部署的 25, 000 多项技能提供了支持。ask 强调第三方开发人员的灵活性、可预测性和快速迭代周期。它带来了归纳偏差, 使其能够从极其小和稀疏的数据集中学习强大的 slu 模型, 并在此过程中消除了软件开发人员和对话系统研究人员进入的重大障碍。少

2018 年 3 月 2 日提交;v1 于 2017 年 11 月 1 日提交;最初宣布 2017 年 11 月。

### 137. 多模式端面效应与抓取系统的设计: 集成设计如何帮助赢得**亚马逊**机器人挑战

作者: s. wade-mccue, n. kelly-boxall, m. mctaggart, d. morrison, a. w. tow, j. erskine, r. Grinover, a. gurman, t. hunn, d. lee, a. 米兰, t. pham, g. rallos, a. razjiaev, t. rowntree, r. smith, k. vijay, z. zan, c. lehnert, i. reid, p. corke j. leitner

**摘要:** 我们提出了掌握系统和设计方法背后的卡特曼, 在 2017 年亚马逊机器人挑战赛的获奖参赛者。我们研究了导致系统最终迭代的设计过程, 并通过与关键机器人设计方面的比较来描述紧急解决方案。根据我们的经验, 我们提出了一个新的设计方面, 精度与冗余, 应该与前面提出的模块化与集成、通用性与假设、计算与实施例和规划的设计方面一起考虑与反馈。我们展示了卡特曼背后的抓取系统, 卡特曼是 2017 年亚马逊机器人挑战赛中获胜的机器人。该系统通过实施免费工具、吸力夹持器和平行夹持器, 在设计中大力利用冗余。这种多模态末端执行器与三种抓取综合算法相结合, 以适应亚马逊在挑战期间提供的对象范围。我们提供了详细的系统描述和对其性能的评估, 然后讨论了系统在机器人设计的关键方面的更广泛性质, 最初是由第一次亚马逊采摘挑战赛的获胜者提出的。为了解决我们抓取系统的主要性质和成功的原因, 我们提出了一个额外的机器人设计方面的 "精度与冗余"。我们的机器人系统的完整设计, 包括末端执行器, 是开源的, 可在 <http://juxi.net/projects/AmazonRoboticsChallenge/> 少

**2018 年 6 月 19 日提交;**v1 于 2017 年 10 月 3 日提交;**最初宣布** 2017 年 10 月。

### 138. 多频抓取和跨域图像匹配的新对象在杂波中的机器人拾取与位置

**作者:**曾志兰,宋树兰,余冠婷, 艾略特·唐龙,弗朗索瓦 r.霍根,玛丽亚·包扎,大林马, 猎户座泰勒, 朱洛迪, 欧德尔·罗莫 nima fazeli, feran alet, nikhil chavan dafle,rachel holladay, isabella morona, prem qu nair, druck green, ian taylor, weber liu, thomas funkhouser, alberto rodriguez

**摘要:**对于已知和新颖的抓取对象, 均可实现较高的精度。该方法是 mit-普林斯顿团队系统的一部分, 在 2017 年亚马逊机器人挑战赛的停滞任务中排名第一。所有代码、数据集和预先培训的模型均可在线获取 <http://arc.cs.princeton.edu>

**2018 年 4 月 1 日提交;**v1 于 2017 年 10 月 3 日提交;**最初宣布** 2017 年 10 月。

### 139. 仓库拾取和放置用笛卡尔机械手的机械设计

**作者:**m. mctaggart, d. morrison, a. w.tow, r.smith, norton kelly-boxall, anton milan , t. pham zheyu zim,j . leitner, i.reid, p. corke, c. lehnert

**摘要:** 机器人操作和在杂乱和非结构化环境中的抓取是当前机器人技术面临的挑战。使机器人能够在这些具有挑战性的环境中运行, 可以直接应用, 从自动化仓库到收获农业水果。与这些困难的机器人操作任务相关的主要挑战之一是多自由度(自由度)机械手的运动规划和控制问题。本文介绍了一个低成本笛卡尔机械手

卡特曼谁在**亚马逊**机器人挑战赛中获得第一名的设计和性能评估。它可以在杂乱无章的环境中执行家庭用品的挑选和放置任务。该机器人的线性速度为1毫米, 角速度为 1.5 rads, 具有亚毫米静态精度和 2kg 的安全有效载荷能力。卡特曼可以生产低于 10, 000 澳元。完整的设计是开源的, 可以在 <http://juxi.net/projects/AmazonRoboticsChallenge> 找到。少

**2018年6月18日提交;**v1 于 2017 年 10 月 2 日提交;**最初宣布** 2017 年 10 月。

#### **140.烟雾筛选器或直射手: 在用户审查社交网络中检测精英 sybil 攻击**

**作者:**郑海忠、薛敏辉、浩路、双浩、朱浩进、梁晓辉、基思·罗斯

**摘要:**流行的用户评论社交网络 (ursn)——如 dianping、yelp 和 amazon——通常是声誉攻击的目标, 在这些攻击中, 为了提高或降低上市产品和服务的评级, 发布虚假评论。这些攻击往往来自于—组名为 sybils 的帐户集合, 这些帐户由—组真正的用户集体管理。一个新的先进计划, 我们称之为精英 sybil 攻击, 招募有机高评级的帐户, 以产生似乎值得信赖和逼真的评论。这些精锐的 sybil 账户综合起来形成了一个庞大的稀疏的 sybil 网络, 现有的 sybil 假审查防御系统不可能成功。在本文中, 我们进行了

第一次研究,以定义、描述和检测西比尔的精英攻击。我们展示了现代精英赛比尔袭击有一个混合架构,第一级招募精锐的西比尔工人,并分配由西比尔组织者的任务,第二层张贴假评论盈利的精英 sybil 工人。我们设计了 ElsieDet,这是一个三阶段的 sybil 检测计划,首先将可疑用户组排除在外,然后确定广告系列窗口,最后确定参加该活动的 sybil 精英用户。我们对中国最受欢迎的 urshn 服务——电平的一千万条评论进行了大规模的实证研究。我们的研究表明,来自优秀的 sybil 用户的评论是更分散的时间,工艺更有说服力的审查,并有更高的过滤旁路率。我们还衡量了 sybil 营销活动对各个行业(如影院、酒店、餐馆)以及连锁店的影响,并证明随着时间的推移监控 sybil 精英用户可以提供针对 sybil 营销活动的宝贵的早期警报。少

2017 年 12 月 4 日提交;v1 于 2017 年 9 月 20 日提交;最初宣布 2017 年 9 月。

#### 141. 卡特曼: 获得亚马逊机器人挑战赛冠军的低成本笛卡尔机械手

作者 :d. morrison, a. w.tow, m.mctaggart, r.smith, n. kelly-boxall, s .Wade-McCue, j. erskine, r. Grinover, a.gurman, t. hunn, d .lee, a. 米兰 , t. pham, g. rallo, a.razjaev, t.rontree, k.vijay, z . zz, c. lehnert, i. reid, p . corke j. leitner

摘要: 亚马逊机器人挑战赛招募了 16 个团队,每个团队设计一个自主仓储的拾取和放置机器人,解决机器人视觉和操作的发展问

题。本文介绍了我们定制的、具有成本效益的笛卡尔机器人系统 cartman 的设计, 该系统在比赛决赛中获得第一名, 在 27 分钟内堆放了 14 个 (16 个项目中的 16 个), 并挑选了所有 9 个项目, 总共得分 272 分。我们强调我们以经验为中心的设计方法和我们系统的关键方面, 这有助于我们的竞争力。我们认为, 这些方面对于建立强大有效的机器人系统至关重要。少

2018 年 2 月 25 日提交;v1 于 2017 年 9 月 19 日提交;最初宣布 2017 年 9 月。

## 142. 深度强化学习聊天机器人

作者 :[iulian v. serban](#), [chinnadhuraisankar](#), [mathieu germain](#), [sazhengzhang](#) , [zh 心 han lin](#), [sandeepsubramanian kim](#), [michael pi 片断 er](#) , [sarath chandar](#) , [nan rosemary ke](#), [sai rajeshwar](#), [亚历山大 debrebisson](#), [jose m. r.sotelo](#), [dendi suhubdy](#), [vincent michalski](#), [亚历山大 nguyen](#), [joelle pineau](#) , [yeshua bengio](#)

**摘要:** 我们介绍 milabot: 一个深入强化学习聊天机器人开发的蒙特利尔学习算法研究所 (mila) 为**亚马逊亚历克莎**奖比赛。milabot 能够通过语音和文本与人类就流行的小谈话话题进行交谈。该系统由自然语言生成和检索模型组成, 包括基于模板的模型、词袋模型、序列序列神经网络和潜在变量神经网络模型。通过将强化学习应用于众包数据和现实世界中的用户交互, 该系

统接受了培训, 以便从其集成中的模型中选择适当的响应。通过对实际用户进行的 aeb 测试对该系统进行了评估, 该系统的性能明显好于许多竞争系统。由于其机器学习架构, 系统很可能会随着额外的数据而改进。少

**2017 年 11 月 5 日提交;****v1** 于 2017 年 9 月 7 日提交;**最初宣布** 2017 年 9 月。

### **143. 基于斑点狮身人面像对齐和特征提取的英语口语智力修复在艺术现状中的显著提高**

**作者:**[yuan gao](#), [brij mohan lal srivastava](#), [james salsman](#)

**摘要:** 基于支持向量机 (svm) 分类器或深度学习神经网络模型确定的学习者口语反应的真实清晰度, 我们使用自动语音识别来评估口语学习者的发音转录正确性的预测。使用齐头并进 etspinx 对齐模式和许多识别通过搜索每个预期的音素和在序列中插入意外音素的数字功能, svm 模型实现了 82% 的一致性。**亚马逊机械土耳其人**群体工作者转录的准确性, 高于多名独立研究人员报告的 75%。利用 svm 分类器概率预测模型的这些功能, 可以帮助计算机辅助语音教学 (capt) 系统提供清晰度修正。少

**2018 年 1 月 26 日提交;****v1** 于 2017 年 9 月 6 日提交;**最初宣布** 2017 年 9 月。



## 144. 推荐人系统中矩阵分解与随机游走与重启的比较研究

作者: [haekyu park](#), [jinhong jung](#), [u kang](#)

**摘要:** 在矩阵分解或随机游走与重新启动 (rwr) 之间, 哪种方法更适合推荐系统? 哪种方法能更好地处理显式或隐式反馈数据? 其他信息是否有助于推荐? 推荐系统在**亚马逊**和网飞等许多电子商务服务中发挥着重要作用, 向用户推荐新的项目。在各种推荐策略中, 通过使用用户的评级模式, 协同过滤表现出了良好的性能。矩阵分解和随机游动是最具代表性的协同过滤方法。然而, 尽管哪种方法具有广泛的效用, 但仍不清楚哪种方法能提供更好的推荐性能。本文对推荐系统中的矩阵分解和 rwr 进行了比较研究。我们根据推荐中的各种任务, 准确地制定了这两种方法的每一个对应项。特别是, 我们新设计了一个使用全局偏置项的 rwr 方法, 该方法对应于一种使用偏置的矩阵分解方法。我们在推荐质量的各个方面描述了这两种方法的细节, 例如这些方法如何处理通常在协同过滤中发生的冷启动问题。我们在实际数据集中进行了大量实验, 以评估每种方法在各种度量方面的性能。我们观察到, 矩阵分解在显反馈额定值下表现更好, 而 rwr 在隐式反馈额定值下表现更好。我们还观察到, 利用项目的全球普及有利于业绩, 侧面信息产生显性反馈的正协同作用, 但对隐含的反馈产生负面影响。少

2017 年 11 月 5 日提交;v1 于 2017 年 8 月 29 日提交;最初宣布 2017 年 8 月。

#### 145. 电子商务加固机构设计

作者:[蔡庆鹏](#),[阿里斯·菲洛斯-拉齐卡斯](#), 唐平忠,[张义伟](#)

**摘要:**我们研究了在亚马逊、易趣或淘宝等电子商务网站上向卖家分配印象的问题,旨在最大限度地提高该平台产生的总收入。采用强化机构设计的总体框架,采用深层加固学习设计高效算法,采取卖家的战略行为。更多

2018 年 2 月 27 日提交;v1 于 2017 年 8 月 24 日提交;最初宣布 2017 年 8 月。

#### 146. 循环 mds 码和扩展图的梯度编码

作者:[netanel raviv](#), [itzhak tamo](#), [r 拉希 s tandon](#), [亚历山大 g. di 马基斯](#)

**摘要:** 梯度编码是分布式学习中一种降低散链器的技术。本文利用经典编码理论中的工具,即循环 mds 码,在参数的适用范围和所涉及算法的复杂性方面,设计了新的梯度码,与现有的解相比是有利的。其次,我们引入了梯度编码问题的近似变量,在这个变量中,我们确定了近似梯度计算,而不是精确的梯度计算。这种方法可以实现优雅的退化,即我 2 近似梯度误差是散向者数量的递

减函数。我们的主要结果是, 扩展器图的归一化邻接矩阵可以产生良好的近似梯度码, 与精确梯度编码相比, 这种方法使我们的计算量明显减少。我们在 **amazon ec2** 上进行了实验测试, 结果表明近似梯度编码的泛化误差非常接近全梯度, 而对工人的计算要求要少得多。少

**2018 年 9 月 18 日提交;v1** 于 2017 年 7 月 12 日提交;**最初宣布** 2017 年 7 月。

#### 147. parlai: 对话研究软件平台

**作者:** [亚历山大·h·米勒](#), [will feng](#), [adam f 午餐](#), [jiasen lu](#), [dhruv batra](#), [antoine bordes](#), [devi parkh](#), [jason weston](#)

**摘要:** 我们介绍了 parlai (发音为 "par-lay"), 这是一个在 python 中实现的对话研究开源软件平台, 可在 <http://parl.ai> 上查阅。其目标是为对话模型的共享、培训和测试、亚马逊机械土耳其人的数据收集、人工评估和在线强化学习提供一个统一的框架;和机器学习模型的存储库, 用于与他人的模型进行比较, 并改进现有的体系结构。第一个版本支持 20 多个任务, 包括热门数据集, 如 squad、babi 任务、mctest、wikqa、qacnn、qadailymail、cbt、babi dialog、ubuntu、open 字幕和 vqa。集成了几种模型, 包括神经网络, 如内存网络、seq2seq 和细心的 lstm。

2018 年 3 月 8 日提交;v1 于 2017 年 5 月 18 日提交;最初宣布 2017 年 5 月。

#### 148. 实时按需人群供电的实体提取

作者:[黄廷浩](#), [陈云农](#), [杰弗里·比格姆](#)

**摘要:** 输出协议机制, 如 esp 游戏已被广泛用于人类计算, 以获得可靠的人为生成的标签。在本文中, 我们认为, 可以使用 "有时限" 的输出协议机制, 在互动系统中创建一个快速而强大的众电组件, 特别是对话系统, 从用户发言中实时提取关键信息。我们使用航空公司旅行信息系统 (atis) 数据集在 amazon 机械土耳其人身上进行的实验表明, 该方法在平均响应时间短于 9 秒的情况下获得高质量的结果。少

2017 年 12 月 6 日提交;v1 于 2017 年 4 月 12 日提交;最初宣布 2017 年 4 月。

#### 149. 小云的云集市动态市场

作者:[ranjan pal](#), [sung-hanlin](#), [aditya ahujay](#), [leana golubchik](#)

**摘要:** ..... 在概念和实践中出现的小云 (sc) 主要是由使用海量数据中心的商业云提供商 (如 amazon) 的服务成本和复杂性相关问题造成的。然而, 由于服务层资源相对稀缺 (如虚拟机), 受

者面临的资源不弹性问题可能会导致服务层的资源不弹性问题。

更多

**2018 年 2 月 13 日提交**;v1 于 2017 年 4 月 3 日提交;**最初宣布** 2017 年 4 月。

## 150. 从子空间构造的数组代码的位置和可用性

**作者:**[natalia silberstein](#), [tuvia etzion](#), [moshe schwartz](#)

**摘要:**在谷歌、脸谱和亚马逊等互联网规模的公司中, 创建和处理的数据数量不断增加。因此, 如此大量数据的有效存储已成为现代计算中的一个根本问题和尖锐问题。没有一台机器能够满足如此巨大的存储需求。因此, 分布式存储系统 (dss)…。更多

**2018 年 6 月 23 日提交**;v1 于 2017 年 1 月 25 日提交;**最初宣布** 2017 年 1 月。

## 151. 利用质量预测模型实现自动写入反馈

**作者:**[hamed Nilforoshan](#), [eugene wu](#)

**摘要:** 用户生成的多段式书写在许多社交媒体平台 (如 amazon 评论、airbnb 主机配置文件等) 中普遍存在且非常重要。确保高质量的内容非常重要。遗憾的是, 用户提交的内容往往不是高质量的。此外, 构成高质量的特征甚至可能因域而异, 用户不知道。

自动写入反馈有可能在编写过程中立即指出并提出改进建议。然而,大多数方法都侧重于语法短语,这只是高质量内容的一个特征。现有的研究开发准确的质量预测模型。我们建议将这些模型与模型解释技术结合起来,以确定如果更改了书写功能,将最大限度地提高文本质量。为此,我们为一类流行的称为树群的模型开发了一种基于摄动的解释方法。此外,我们使用弱监控技术来调整这种方法,以生成特定文本段的反馈,以及整个文档的反馈。我们的用户研究发现,基于扰动的方法,如果与特定于段的反馈相结合,可以帮助提高 **amazon** (审查有用性) 和 **airbnb** (主机配置文件可信度) 上的写作质量 > 14% (3x 改善最近的自动反馈技术)。少

**2018 年 4 月 18 日提交**;v1 于 2017 年 1 月 23 日提交;**最初宣布** 2017 年 1 月。

## 152. 异构聚类的编码计算

**作者** :[amirhoshussein reisizadeh](#), [saurav prakash](#), [ramtin pedarsani](#), [amir salman avestimehr](#)

**摘要**: 在大规模分布式计算集群 (如 **amazon ec2**) 中,有几种类型的 "系统噪声" 可能会导致性能严重下降: 通信带宽有限导致的瓶颈、由散点节点引起的延迟、等。另一方面,这些系统具有丰富的冗余——大量的计算节点和大的存储容量。最近的研究结果表明,编码对于有效利用计算和存储冗余以缓解同质集群中的散向

标和通信瓶颈的影响具有重要意义。本文重点研究了由各种具有不同功能的计算机组成的一般异构分布式计算集群。我们提出了一个编码框架, 通过交易冗余来减少计算延迟, 从而加快异构集群中的分布式计算。特别是, 我们提出了异构编码矩阵乘法 (hcmm) 算法, 用于在异构群上执行分布式矩阵乘法, 对于广泛的处理时间分布, 异构聚类是可实现的渐近最优算法。此外, 我们还表明, hcmm 比任何未编码的方案都要快得多。为了证明 hcmm 的实用性, 我们在 **amazon ec2** 集群上进行了实验, 发现 hcmm 达到了 61%, 46% 和 36% 分别比三个基准负载分配方案更快—统一未编码、负载平衡未编码和统一编码。此外, 我们还对异构设置中的最优负载分配问题进行了概括, 其中我们考虑到了与集群相关的货币成本。我们认为 hcmm 在预算约束的场景中也是渐近最优的, 我们为预算有限的计算任务开发了一种计算 (hcmm) 负载分配的启发式算法。少

**2018 年 10 月 9 日提交;**v1 于 2017 年 1 月 20 日提交;**最初宣布** 2017 年 1 月。

### 153. 恶意事件分类语义关键字的开发

**作者 :** [hyungtaelee](#), [sungmin eum](#), [joel levis](#), [hesung kwon](#), [james michaelis](#), [michael kolodny](#)

**摘要:** 当场景在语义上不同但视觉上相似时, 学习事件分类器是很有挑战性的。然而, 作为人类, 我们通常通过添加背景语义知识

来无痛地处理此类任务。在这一观察的推动下，我们的目标是提供一个实证研究，说明语义关键字等额外信息如何能促进对此类事件的歧视。为了证明这项研究的有效性，我们首先构建了一个新的恶意人群数据集，其中包含两个事件，良性和恶意，在视觉上看起来相似。请注意，本文的主要重点不是在此数据集上提供最先进的性能，而是显示使用语义驱动的关键字信息的有益方面。通过利用众包平台（如 **amazon** 之所以风云机，我们收集与图像相关的语义关键字，然后识别特定事件特有的关键字子集（如警察、消防等）。我们首先表明，通过使用最近引入的注意力模型，一个天真的基于 **cnn** 的事件分类器实际上学习了主要关注与土耳其人识别的判别语义关键字相关的局部属性。我们进一步表明，将关键字驱动的信息集成到早期和后期的融合方法中，可以显著增强恶意事件分类。少

**2017 年 12 月 1 日提交**;v1 于 2016 年 10 月 21 日提交;**最初宣布** 2016 年 10 月。

#### 154. 推荐人系统的自适应非负矩阵分解与测量比较

**作者:**[gianna m. del corso](#), [francesco romani](#)

**摘要:** 评级矩阵的非负矩阵分解 (nmf) 已被证明是解决推荐问题的有效方法。本文提出了一种基于评级矩阵 nmf 的新方法，并将其与 svd 和正正化、非正则化的非负矩阵分解方法等经典算法进



行了比较。特别是得到了一种新的算法, 该算法自适应地改变了每个步骤要最小化的函数, 实现了一种动态的先验策略。通过强制对上一个期限的未知评级进行重构, 得到了另一种算法, 从而修改了 nmf 公式中要最小化的函数。然后, 我们结合不同的方法, 获得两种混合策略, 结果是非常有效的重建缺失的观测。我们在若干评价措施的基础上, 对不同的方法进行了深思熟虑的比较。我们特别考虑评级、分类和排名指标, 表明在考虑不同的措施时, 获得给定度量值的算法总体上也是最好的, 从而降低了设计特定措施的兴趣评估措施。这些算法已在不同的数据集 (特别是 1m) 和 10m movielens 数据集上进行了测试, 这些数据集包含电影收视率、jester 数据集 (在笑话上大发雷霆) 和亚马逊精细食品数据集 (amazon fine food 数据集) 对食品进行了评级。通过对不同算法的比较, 表明了采用显式和隐式正则化方案的方法的良好性能。此外, 我们可以通过将快速方法与更准确的方法相结合的混合策略得到提升。少

**Submitted** 27 August, 2018; **v1submitted** 26 July, 2016; **originally announced** July 2016.