

[BT](#)

- [投稿](#)
- [关于我们](#)
- [合作伙伴](#)

• 欢迎关注我们的：

- 
- 
- 

InfoQ – 促进软件开发领域知识与创新的传播



[登录](#)



- [En](#) |
- [中文](#) |
- [日本](#) |
- [Fr](#) |
- [Br](#)

482,381 十一月 独立访问用户

- [语言 & 开发](#)
 - [Java](#)
 - [.Net](#)
 - [云计算](#)
 - [移动](#)
 - [HTML 5](#)
 - [JavaScript](#)
 - [Ruby](#)
 - [DSLs](#)
 - [Python](#)
 - [PHP](#)
 - [PaaS](#)

特别专题语言 & 开发

[技术人该做什么样的产品？](#)



创业两次，仍未成功，一直在迷惘中前行，不断地反思，还没有结果。心中有一个疑问想分享给大家，一个热爱技术的人，该做什么一个产品，该怎么做一个产品，用我自己的经历试图回答。

[浏览所有语言 & 开发](#)

- [架构 & 设计](#)
 - [建模](#)
 - [性能和可伸缩性](#)
 - [领域驱动设计](#)
 - [AOP](#)
 - [设计模式](#)
 - [安全](#)
 - [云计算](#)
 - [SOA](#)

特别专题架构 & 设计

[技术人该做什么样的产品？](#)



创业两次，仍未成功，一直在迷惘中前行，不断地反思，还没有结果。心中有一个疑问想分享给大家，一个热爱技术的人，该做什么一个产品，该怎么做一个产品，用我自己的经历试图回答。

[浏览所有架构 & 设计](#)

- [过程 & 实践](#)

InfoQ读者们...【ArchSummit北京2014】 将于12月19-20日举行，这是ArchSummit首次来到北京，我们将提供大会同步直播

12月19日上午9:30开始，将带来3场主题演讲直播；19日13:30-17:30将对【疯狂双十一】专题进行直播。20日9:30-17:30将对【移动互联网】专题进行直播。

特别专题过程 & 实践

[不只是LeSS](#)



虽然敏捷社区对大规模敏捷的实施带来了令人耳目一新的方法，但围绕大项目来解决组织的复杂性问题，这些模型似乎仍然有些匮乏。本文对大规模Scrum提供了一个全面的方法。它是在LeSS的基础上做了一定的修正，以便更好地面对大型项目的挑战。

[浏览所有过程 & 实践](#)

- [运维 & 基础架构](#)
 - [性能和可伸缩性](#)
 - [大数据](#)
 - [DevOps](#)
 - [云计算](#)
 - [虚拟化](#)
 - [NoSQL](#)
 - [应用服务器](#)
 - [运维](#)

特别专题运维 & 基础架构

[从MDM到MDM：从管理设备到管理应用再到管理数据](#)



移动业务管理（Mobile operations management, MOM）如今对于现代化的企业来讲是具有顶级优先级的事情。第一代的MOM解决方案围绕着移动设备和管理，组织机构很快意识到有效的企业级移动基础设施需要解决一个更为困难的问题：管理和保护

移动业务数据。

[浏览所有运维 & 基础架构](#)

- [企业架构](#)
 - [企业架构](#)
 - [业务流程建模](#)
 - [业务/IT整合](#)
 - [Integration \(EAI\)](#)
 - [治理](#)
 - [Web 2.0](#)
 - [SOA](#)

特别专题企业架构

[从MDM到MDM：从管理设备到管理应用再到管理数据](#)



移动业务管理（Mobile operations management，MOM）如今对于现代化的企业来讲是具有顶级优先级的事情。第一代的MOM解决方案围绕着移动设备和管理，组织机构很快意识到有效的企业级移动基础设施需要解决一个更为困难的问题：管理和保护移动业务数据。

[浏览所有企业架构](#)

ArchSummit 全球架构师峰会

北京 | 2014年12月19-20日

QCon 全球软件开发大会

北京 | 2015年4月23-25日

[移动](#)

[Docker](#)

[Node.js](#)

[云计算](#)

[大数据](#)

[架构师](#)

[ArchSummit](#)

[AWS](#)

[Azure](#)

[全部话题](#)

机器学习的11个开源项目

作者 [张天雷](#) 发布于 2014年12月18日 / [讨论](#)

分享到: [微博](#) [微信](#) [Facebook](#) [Twitter](#) [有道云笔记](#) [邮件分享](#)

- [“稍后阅读”](#)
- [“我的阅读清单”](#)

机器学习是目前数据分析领域的一个热点内容，在平时的学习和生活中经常会用到各种各样的机器学习算法。实际上，基于Python、Java等的很多机器学习算法基本都被前人实现过很多次了。这些算法在网上可以找到很多，然而往往存在很多“脏”或者“乱”的开源代码。

在这样的背景下，[InfoWorld](#)近日公布了机器学习领域11个最受欢迎的开源项目，这11个开源项目大多与垃圾邮件过滤、人脸识别、推荐引擎相关。它们大多数基于现今最流行的语言以及平台，推广以及扩展了机器学习领域的很多重要算法。从中，用户不但可以找到LDA等主题模型，也可以找到HMM等隐马尔科夫模型。这些模型都是应用领域的热点，也是研究者们最需要的。

1. [Scikit-learn](#)

Scikit-learn是一个非常强大的Python机器学习工具包。它通过在现有Python的基础上构建了NumPy和Matplotlib，提供了非常便利的数学工具。这个工具包包括了很多简单且高效的工具，很适合用于数据挖掘和数据分析。

在主页中，可以看到User Guide，这是整个机器学习的索引，其中用户可以学到各种有效的方法。在Reference里，用户可以找到各个类具体的用法索引。

2. [Shogun](#)

Shogun是一个基于C++的最古老的机器学习开源库，它创建于1999年。作为一个SWIG库，Shogun可以轻松地嵌入Java、Python、C#等主流处理语言中。它的重点在于大尺度上的内核方法，特别是“支持向量机”的学习工具箱。其中，它包括了大量的线性方法，如LDA、LPM、HMM等等。

3. [Accord Framework/AForge.net](#)

Accord是AForge.net的扩展，是一个基于.Net的机器学习与信号处理框架。它包括了一系列的对图像和音频的机器学习算法，如人脸检测、SIFT拼接等等。同时，Accord支持移动对象的实时跟踪等功能。它提供了一个从神经网络到决策树系统的机器学习库。

4. [Mahout](#)

Mahout是一个广为人知的开源项目，它是Apache Software旗下的一个开源项目，提供了众多的机器学习经典算法的实现，旨在帮助开发人员更加方便快捷地创建智能应用程序。Mahout内包含了聚类、分类、推荐等很多经典算法，并且提供了很方便的云服务的接口。

5. [MLlib](#)

MLlib是Apache自己的Spark和Hadoop机器学习库，它被设计用于大规模高速度地执行MLlib所包含的大部分常见机器学习算法。MLlib是基于Java开发的项目，同时可以方便地与Python等语言对接。用户可以自己设计针对MLlib编写代码，这是很具有个性化的设计。

6. [H2O](#)

H2O是0xdata的旗舰产品，是一款核心数据分析平台。它的一部分是由R语言编写的，另一部

分是由Java和Python语言编写的。用户可以部署H2O的R程序安装包，之后就可以在R语言环境下运行了。H2P的算法是面向业务欺诈趋势预测的，目前正在新一轮的融资中。

7. [Cloudera Oryx](#)

Oryx也是由Hadoop所设计的机器学习开源项目，由Cloudera Hadoop Distribution的创造者所提供。Oryx能够让机器学习的模型使用在实时的数据流上，如垃圾邮件过滤等。

8. [GoLearn](#)

GoLearn是谷歌所构建的Go语言的一体化机器学习库，目标是简单并且可定制。Go语言是谷歌的主打语言，目前使用已经越来越广泛。GoLearn的简单在于数据在库内被加载和处理，因此能够可定制地扩展数据结构以源码。

9. [Weka](#)

>Weka是使用Java开发的用于数据挖掘的开源项目。Weka作为一个公开的数据挖掘工作平台，集合了大量能够承担数据挖掘任务的机器学习算法，包括了对数据进行预处理、分类、回归、聚类等等。同时，Weka实现了对大数据的可视化，通过Java设计的新式交互界面上，实现人与程序的交互。

10. [CUDA-Convnet](#)

CUDA是我们众所周知的GPU加速套件。而CUDA-Convnet是一个基于GPU加速的神经网络应用程序机器学习库。它使用C++编写，并且使用了NVIDIA的CUDA GPU处理技术。

目前，这个项目已经被重组成为CUDA-Convnet2，支持多个GPU和Kepler-generation GPUs。Vuples项目与之类似，使用F#语言编写，并且适用于.Net平台上。

11. [ConvNetJS](#)

ConvNetJS是一款基于JavaScript的在线深度学习库，它提供了在线的深度学习训练方式。它能够帮助深度学习的初学者更快、更加直观的理解算法，通过一些简单的Demo给用户最直观的解释。

感谢[郭莹](#)对本文的审校。

给InfoQ中文站投稿或者参与内容翻译工作，请邮件至editors@cn.infoq.com。也欢迎大家通过新浪微博（@InfoQ）或者腾讯微博（@InfoQ）关注我们，并与我们的编辑和其他读者朋友交流。

【[ArchSummit北京2014](#)】解读“双11”背后淘宝、京东、小米的技术支撑。年末的技术盛宴，CTO与架构师的大Party，[了解详情](#)。

- [领域](#)
- [架构 & 设计](#)
- [语言 & 开发](#)
- [专栏](#)
- [开放源代码](#)
- [开源项目发布](#)
- [机器学习](#)
- [开源软件](#)
- [开源](#)

相关内容

[Google Cardboard: 来自谷歌的虚拟现实开源项目](#)

[章文嵩：怎样做开源才有意义？](#)

[拥抱开源的微软](#)

[拥抱开源：小米的经验分享](#)

[Github开源其文本编辑器Atom](#)

相关厂商内容

[“双11”背后淘宝、京东、一号店的技术支撑](#)

[Windows Azure从开发到部署的自动化进程](#)

[微软Windows Azure开启机器学习之旅](#)

[基于开源软件的Azure平台大规模系统构建](#)

[AWS中众多存储选项、理想使用情况及弹性和可扩展性等特性。](#)

相关赞助商

Windows Azure专区上线，全面了解云服务[精彩呈现！](#)



您好，朋友！

您需要 [注册一个InfoQ账号](#) 或者 [登录](#) 才能进行评论。在您完成注册后还需要进行一些设置。

获得来自InfoQ的更多体验。

告诉我们您的想法

| | |
|-------|----|
| 请输入主题 | 信息 |
|-------|----|

允许的HTML标签：a, b, br, blockquote, i, li, pre, u, ul, p

☐ 当有人回复此评论时请E-mail通知我

社区评论 [Watch Thread](#)
[关闭](#)

by

发布于

- [查看](#)
- [回复](#)
- [回到顶部](#)

[关闭](#)

主题 您的回复

允许的HTML标签：a, b, br, blockquote, i, li, pre, u, ul, p

☐ 当有人回复此评论时请E-mail通知我

[关闭](#)

主题 您的回复

允许的HTML标签：a, b, br, blockquote, i, li, pre, u, ul, p

☐ 当有人回复此评论时请E-mail通知我



[关闭](#)

深度内容

- [全部](#)
- [文章](#)
- [演讲](#)
- [访谈](#)
- [迷你书](#)

[技术人该做什么样的产品？](#)

[郝培强](#) 12月18日



[GS Collections实例教程（第一部分）](#)

[Donald Raab](#) 12月18日





[智慧风能与云计算](#)

[潘怡鸿](#) 12月17日



[Spark应用案例分析](#)

[田毅](#) 12月17日



[如何评判产品特性的“价值”](#)

[姚安峰](#) 12月17日



[Hibernate资深开发者刘少壮谈创业心得：团队凝聚力首先来自招聘爱干这个事情的人](#)

[刘少壮](#) 12月16日



- [更早的 >](#)

赞助商链接

以智能硬件为代表的物联网产品引领新一波的创业浪潮，开放物联网技术大会带您了解物联网行业的发展趋势及相关的创业项目。 [马上报名!](#)

【云平台技术全景剖析】12月19日，听黄湘龙（龙觉）、张献涛(旭卿)、朱照远（叔度）、占超群(离哲)等技术专家分享云平台背后的核心技术。 [现在报名!](#)

Windows Azure系列课程将推出 Microsoft Azure 在计算 Compute 服务上的所有功能，包括虚拟机、云服务、网站和移动服务等。更多分享请锁定Windows Azure专区。 [点击进入>>](#)
语言 & 开发

[年末技术盛宴ArchSummit北京2014大会举行](#)

[技术人该做什么样的产品？](#)

[如何把iOS代码编译为Android应用 - iOS移动开发周报](#)

架构 & 设计

[年末技术盛宴ArchSummit北京2014大会举行](#)

[技术人该做什么样的产品？](#)

[如何把iOS代码编译为Android应用 - iOS移动开发周报](#)

过程 & 实践

[开放、开源、一站式：首届开放物联网大会（OIOT）在京举办](#)

[初用敏捷：必须从组织架构入手吗？](#)

[不只是LeSS](#)

运维 & 基础架构

[开放、开源、一站式：首届开放物联网大会（OIOT）在京举办](#)

[Mahout通过可插拔的后端平台Spark和Flink获取自优化矩阵代数接口](#)

[Plumbr转型为JVM监控解决方案](#)

企业架构

[开放、开源、一站式：首届开放物联网大会（OIOT）在京举办](#)

[初用敏捷：必须从组织架构入手吗？](#)

[Typesafe调查：Java 8广泛普及，用户热切盼望Java 9](#)

- [首页](#)
- [全部话题](#)
- [QCon全球软件开发大会](#)
- [关于我们](#)
- [投稿](#)
- [创建账号](#)
- [登录](#)
- [全球QCon](#)
- [伦敦 2015年3月2-6日](#)
- [圣保罗 2015年3月23-27日](#)
- [北京 2015年4月23-25日](#)
- [东京 2015年4月 21](#)
- [纽约 2015年6月8-12日](#)
- [里约 2015年8月24-28日](#)
- [上海 2015年10月15-17日](#)
- [旧金山 2015年11月16-20日](#)

InfoQ每周精要

通过个性化定制的新闻邮件、RSS Feeds和InfoQ业界邮件通知，保持您对感兴趣的社区内容的时刻关注。

[点击这里
查看样刊](#)



- [属于您的个性化RSS](#)
- [InfoQ官方微博](#)
- [InfoQ官方微信](#)
- [社区新闻和热点](#)

特别专题

- [IBM](#)
- [技术社区活动日历](#)
- [百度技术沙龙](#)

- [月刊：《架构师》](#)
- [线下活动：QClub](#)
- [AWS专区](#)
- [物联网大会](#)
- [Code Rally 编程拉力赛](#)

定制您感兴趣的技术领域

- ☒ 语言 & 开发
- ☒ 架构 & 设计
- ☒ 过程 & 实践
- ☒ 运维 & 基础架构
- ☒ 企业架构

这会影响您在主页和RSS订阅中看到的内容。点击“偏好设置”可选择更多精彩定制内容。

提供反馈

错误报告

商务合作

内容合作

feedback@cn.infoq.com bugs@cn.infoq.com sales@cn.infoq.com editors@cn.infoq.com

InfoQ.com
及所有内
容，版权所
有 © 2006-
2014
C4Media
Inc.
InfoQ.com
服务器由
[Contegix](#)提
供，我们最
信赖的ISP
伙伴。
北京创新网
媒广告有限
公司 京ICP
备09022563
号-7 [隐私
政策](#)

[BT](#)