

From GITHUB [awesome-machine-learning](https://github.com/josephmislove/awesome-machine-learning)

机器学习相关资料目录

目录

目录

- [C](#)
- [通用机器学习](#)
- [计算机视觉](#)
- [C++](#)
- [计算机视觉](#)
- [通用机器学习](#)
- [自然语言处理](#)
- [序列分析](#)
- [手势识别](#)
- [Common Lisp](#)
- [通用机器学习](#)
- [Clojure](#)
- [自然语言处理](#)
- [通用机器学习](#)
- [数据分析 / 数据可视化](#)
- [Erlang](#)
- [通用机器学习](#)
- [Go](#)
- [自然语言处理](#)
- [通用机器学习](#)
- [数据分析 / 数据可视化](#)
- [Haskell](#)
- [通用机器学习](#)
- [Java](#)
- [自然语言处理](#)
- [通用机器学习](#)
- [数据分析 / 数据可视化](#)
- [深度学习](#)

- [Javascript](#)
- [自然语言处理](#)
- [数据分析 / 数据可视化](#)
- [通用机器学习](#)
- [其他](#)
- [Julia](#)
- [通用机器学习](#)
- [自然语言处理](#)
- [数据分析/数据可视化](#)
- [其他项 / 展示](#)

C

通用机器学习

- [Recommender](#) - 一个C语言库，利用协同过滤（CF）进行产品推荐/建议；
- [Darknet](#) - 是一个用C和CUDA编写的开源神经网络框架，它速度快，易于安装，并支持CPU和GPU计算。

计算机视觉

- [CCV](#) - 基于C语言、高速缓存的核计算机视觉库，是一个现代化的计算机视觉库；
- [VLFeat](#) - 开放、可便携的计算机视觉算法库，内有matlab工具箱。

语音识别

- [HTK](#) - 隐马尔可夫模型工具包（HTK）是一个便携式工具包，用于构建和操作隐马尔可夫模型。

C++

计算机视觉

- [OpenCV](#) - OpenCV自带C ++、C、Python、Java和MATLAB接口，并支持Windows、Linux、Android版和Mac OS等系统；
- [DLib](#) - DLib有C ++和Python接口，用于人脸检测和训练通用的目标探测器；
- [EBLearn](#) - Eblearn是一种面向对象的C++库，能够实现各种机器学习模型；
- [VIGRA](#) - VIGRA是通用跨平台的C++计算机视觉和机器学习库，能够用Python绑定任意维度的体积。

通用机器学习

- [mlpack](#) - 可扩展的C++机器学习库；
- [DLib](#) - ML工具套件，能够很容易嵌入到其他应用程序中；
- [encog-cpp](#)
- [shark](#)
- [Vowpal Wabbit \(VW\)](#) - 一个快速的核外（out-of-core）学习系统；

- [sofia-ml](#) -快速增量算法套件；
- [Shogun](#) -Shogun机器学习工具箱；
- [Caffe](#) -基于清洁度、可读性和速度考虑而开发的深度学习框架。[深度学习]
- [CXXNET](#) -另一个深度学习框架，其核心代码少于1000行；[深度学习]
- [XGBoost](#) -一种并行、优化、通用的梯度推进库；
- [CUDA](#) -利用C++ / CUDA快速实现卷积；[深度学习]
- [Stan](#) -一种概率性的编程语言，能够实现Hamiltonian Monte Carlo抽样的全贝叶斯统计推断；
- [BanditLib](#) -一个简单、有多重保护的Bandit库；
- [Timbl](#) : 一个软件包/ C++库，能够实现多种基于内存的学习算法，其中有IB1-IG--k-最近邻分类的实现、IGTree--IB1-IG的决策树近似值；常用于NLP；
- [Disrtibuted Machine learning Tool Kit \(DMTK\)](#) -微软开发的分布式机器学习（参数服务器）框架，能够在多台机器的大型数据集上实现训练模型，与它捆绑的现有工具包括：LightLDA和分布式（多传感）字嵌入；
- [igraph](#) -通用图形库；
- [Warp-CTC](#) -在CPU和GPU上快速并行地实现连接时域分类（Connectionist Temporal Classification, CTC）；
- [CNTK](#) -微软研究院开发的计算网络工具包（CNTK），它作为一系列计算步骤，通过有向图来描述神经网络，是统一的深度学习工具包；
- [DeepDetect](#) -一个机器学习API，服务器用C++11编写，它使机器学习的状态易于工作，并容易集成到现有应用程序；
- [Fido](#) -一个高度模块化的C++机器学习库，用于嵌入式电子产品和机器人中。

自然语言处理

- [MIT Information Extraction Toolkit](#) -C, C++和Python的工具，用于命名实体识别与关系抽取；
- [CRF++](#) -条件随机域（Conditional Random Fields, CRFs）的开源实现，用于分割/标记序列数据及其他自然语言处理任务；
- [CRFsuite](#) -条件随机域（CRFs）的实现，用于标记序列数据；
- [BLLIP Parser](#) -BLLIP自然语言解析器（也称为Charniak-Johnson解析器）；
- [colibri-core](#) -是C++库、命令行工具，和Python绑定用于提取与使用基本的语言结构，例如用快速和高效存储的方式实现n-grams和skipgrams模型；
- [ucto](#) -是一种工具和C++库，基于支持各种语言的编译器，内含统一字符标准及规则表达式；支持FoLiA格式；
- [libfolia](#) -支持FoLiA格式的C++库；

- [frog](#) - 为Dutch开发的基于内存的NLP套件：POS标签、归类分析、依存句法分析、NER、浅层句法分析、形态分析；
- [MeTA](#) - [MeTA : ModErn 语篇分析 \(ModErn Text Analysis\)](#)，是一个C++数据科学工具包，便于挖掘大文本数据。

语音识别

- [Kaldi](#) - Kaldi是用于语音识别的工具包，用C++编写，由Apache许可证V2.0协议授权，专门给语音识别的研究人员使用。

序列分析

- [ToPS](#) - 这是一种面向对象的框架，便于在用户定义的字母序列中整合概率模型。

手势检测

- [grt](#) - 手势识别工具包（GRT）是一个跨平台、开源的C++机器学习库，用于实时的手势识别。

Common Lisp

通用机器学习

- [mgl](#) - 神经网络（玻耳兹曼机，前馈和循环网络）以及高斯过程；
- [mgl-gpr](#) - 演化算法；
- [cl-libsvm](#) - LIBSVM支持向量机库的包装。

Clojure

自然语言处理

- [Clojure-openNLP](#) - Clojure中自然语言处理的工具包（openNLP）；
- [Infections-clj](#) - Clojure和Clojure中类似于Rails的inflection库。

通用机器学习

- [Touchstone](#) - Clojure A/B 测试库；
- [Clojush](#) - Clojure中的Push程序语言和PushGP遗传编程系统；
- [Infer](#) - Clojure中分析和机器学习的工具；
- [Cli-ML](#) - Clojure中基于Weka及其相关环境的深度学习库；
- [Encog](#) - 在Clojure中封装成Encog（v3）（专门研究神经网络的机器学习框架）；
- [Fungp](#) - Clojure中的遗传编程实例库；
- [Statistiker](#) - Clojure中基础机器学习算法；
- [clortex](#) - 采用Numenta's Cortical 学习算法的通用机器学习库；
- [comportex](#) - 采用Numenta's Cortical 学习算法的功能组合的机器学习库。

数据分析/数据可视化

- [Incanter](#) - Incanter是基于 Clojure，类似R的统计计算与制图平台；
- [PigPen](#) - Clojure中的Map-Reduce；

- [Envision](#) - 基于Statistiker和D3Clojure 数据可视化库。

Erlang

通用机器学习

- [Disco](#) - Erlang中的Map Reduce模型。

Go

自然语言处理

- [go-porterstemmer](#) - 一个用于实现Porter词干提取算法的原生Go语言净室；
- [paicehusk](#) - Go语言中用于实现Paice/Husk词干提取算法；
- [snowball](#) - Go语言中的Snowball 词干提取器；
- [go-ngram](#) -内存N-gram索引压缩 。

通用机器学习

- [Go Learn](#) - Go语言中的机器学习库；
- [go-pr](#) - Go语言中的模式识别包；
- [go-ml](#) - 线性/逻辑回归、神经网络、协同过滤和多元高斯分布；
- [bayesian](#) - Go语言中朴素贝叶斯分类库；
- [go-galib](#) - Go语言版的遗传算法库；
- [Cloudforest](#) - GO语言中的决策树集合；
- [gobrain](#) - GO语言版的神经网络；
- [GoNN](#) - GoNN 是用Go语言实现的神经网络，它包括BPNN、RBF、PCN ；
- [MXNet](#) - 轻量级、便携式、灵活的分布式/深度学习系统，可对动态的、突变数据流调度部署，同时也支持Python、R、Julia、Go、Javascript 等编程语言。

数据分析/数据可视化

- [go-graph](#) - Go语言图形库；
- [SVGo](#) - Go语言的SVG生成库；
- [RF](#) - Go语言的随机森林库；

Haskell

通用机器学习

- [haskell-ml](#) - Haskell 语言实现的各种深度学习算法 ；
- [HLearn](#) - 根据代数结构解释其深度模型的库；
- [hnn](#) - Haskell语言的神经网络库；
- [hopfield-networks](#) - Haskell中用于无监督学习的Hopfield网络；
- [caffegraph](#) - 一种用于深度神经网络的领域特定语言（DSL）；

- [LambdaNet](#) - Haskell中的可配置的神经网络。

Java

自然语言处理

- [Cortical.io](#) - 像人脑一样快速、精确处理复杂的NLP（自然语言处理）操作（如消歧、分类、流文本过滤等操作）的Retina API；
- [CoreNLP](#) - 斯坦福大学的CoreNLP提供的一系列的自然语言处理工具，该工具可以根据输入原始英语文本，给出单词的基本形式；
- [Stanford Parser](#) - 一种自然语言分析器，可以分析语句的语法结构；
- [Stanford POS Tagger](#) - 一个词性分类器（POS Tagger）；
- [Stanford Name Entity Recognizer](#) - Stanford NER是一个Java实现的名词识别器；
- [Stanford Word Segmenter](#) - 分词器，很多NLP工作中都要用到的标准预处理步骤；
- [Tregex, Tsurgeon and Sengrex](#) - Tregex基于树关系以及节点匹配的正则表达式，用于在树状数据结构中进行模式匹配（名字是“tree regular expressions”的缩写）；
- [Stanford Phrasal](#): 一个基于短语的翻译系统
- [Stanford English Tokenizer](#) - Stanford Phrasal 用Java写成的最新的基于统计短语的机器翻译系统；
- [Stanford Tokens Regex](#) - 一个分解器，可以将文本大致分成一系列对应于“词”的符号；
- [Stanford Temporal Tagger](#) - SUTime是一个用于识别并标准化时间表达式的库；
- [Stanford SPIED](#) - 在种子集上使用模式，以迭代方式从无标签文本中进行学习实体；
- [Stanford Topic Modeling Toolbox](#) - 为社会科学家及其他希望分析数据集的人员提供的主题建模工具；
- [Twitter Text Java](#) - Java实现的推特文本处理库；
- [MALLET](#) - 基于Java的统计自然语言处理、文档分类、聚类、主题建模、信息提取以及其他机器学习的文本应用包；
- [OpenNLP](#) - 基于机器学习的自然语言文本处理工具包；
- [LingPipe](#) - 一个使用计算机语言学文本处理的工具包；
- [ClearTK](#) - 在Java中ClearTK为开发统计语言处理组件提供了一个框架，该框架是基于Apache UIMA；
- [Apache cTAKES](#) - Apache cTAKES是一个开源自然语言处理系统，用于从临床电子病历的自由文本中提取信息；
- [ClearNLP](#) - ClearNLP工程为自然语言处理提供了软件和资源提供了。该项目最早在计算机愈合和教育研究中心启动，目前由Emory 大学的语言和信息研究中心继续开发。该项目遵循Apache 2 license。

通用机器学习

- [aerosolve](#) - 是由Airbnb设计的定价建议系统的机器学习库；
- [Datumbox](#) - 应对机器学习和统计应用快速发展的机器学习框架；
- [ELKI](#) - 用于数据挖掘的Java工具包(无监督：聚类、异常检测等等)；
- [Encog](#) - 一种先进的神经网络和机器学习框架。 Encog包含用于创建各类网络的类，同时也支持为神经网络规范和处理数据的类。它的训练采用多线程弹性传播。它也能使用GPU加快处理时间。提供了图形化界面来帮助建模和训练神经网络。
- [EvA2](#) - 包含遗传算法、差分进化、协方差自适应进化策略等等的进化算法框架；
- [FlinkML in Apache Flink](#) - Flink中的分布式机器学习库；
- [H2O](#) - 深度学习引擎，支持在Hadoop、Spark 或者通过R、Python、Scala 、REST/JSONML 的APIs连到的笔记本上进行分布式学习；
- [htm.java](#) - 采用Numenta Cortical 学习算法的通用机器学习库 ；
- [java-deeplearning](#) - Java、Clojure、Scala的分布式深度学习平台；
- [JAVA-ML](#) - 包含所有Java算法的通用接口的通用深度学习库；
- [JSAT](#) - 用于分类、回归、聚类的机器学习算法集合；
- [Mahout](#) - 分布式的机器学习库；
- [Meka](#) - MEKA提供了一个面向多标签学习和评价方法的开源实现（扩展成Weka）；
- [MLlib in Apache Spark](#) - Spark中的分布式机器学习程序库；
- [Neuroph](#) - Neuroph 是轻量级的Java神经网络框架；
- [ORYX](#) - 采用Apache Spark和Apache Kafka的Lambda 结构框架，专门用于实时大规模机器学习；
- [Samoa](#) - SAMOA 是一个包含用于分布式机器学习数据流的框架，同时为数据流流入不同的流处理平台提供了接口；
- [RankLib](#) - RankLib是一个排序学习算法库；
- [rapaio](#) - Java中用于统计、数据挖掘和机器学习的工具箱；
- [RapidMiner](#) - RapidMiner integration into Java code
- [Stanford Classifier](#) - 斯坦福大学分类器是一种机器学习工具，它可以将数据项归置不同的类别中；
- [SmileMiner](#) - 统计机器智能与学习引擎；
- [SystemML](#) - 灵活、可扩展的机器学习语言；
- [WalnutiQ](#) - 人脑部分面向对象模型；
- [Weka](#) - Weka是数据挖掘方面的机器学习算法集 。

语音识别

- [CMU Sphinx](#) - CMU Sphinx 是基于Java 语音识别库，用于纯语音识别开源工具包。

数据分析/数据可视化

- [Flink](#) - Apache Flink是一个面向分布式数据流处理和批量数据处理的开源计算平台;
- [Hadoop](#) - 大数据分析平台;
- [Spark](#) - Spark是一个快速通用的大规模数据处理引擎;
- [Storm](#) - Storm是一个分布式实时计算系统;
- [Impala](#) - 为Hadoop实现实时查询
- [DataMelt](#) - 用于数字计算、统计、符号计算、数据分析和数据可视化的数学软件;
- [Dr. Michael Thomas Flanagan's Java Scientific Library](#)

深度学习

- [Deeplearning4j](#) - 采用并行GPU的商用可扩展深度学习库。

数 Javascript

自然语言处理

- [Twitter-text](#) - Twitter文本处理库中使用JavaScript的实现;
- [NLP.js](#) - 使用JavaScript和CoffeeScript的NLP实用工具;
- [natural](#) - 用于节点的通用自然语言工具;
- [Knwl.js](#) - JS中的自然语言处理器;
- [Retext](#) -用于分析和处理自然语言的可扩展系统;
- [TextProcessing](#) -情感分析, 词干和词形还原, 部分词性标注和组块, 短语提取和命名实体识别;
- [NLP Compromise](#) - 浏览器中的自然语言处理。

数据分析/数据可视化

- [D3.js](#)
- [High Charts](#)
- [NVD3.js](#)
- [dc.js](#)
- [chart.js](#)
- [dimple](#)
- [amCharts](#)
- [D3xter](#) - 直接建立在D3上的绘图;
- [statkit](#) - JavaScript的统计工具;
- [datakit](#) - JavaScript的轻量级数据分析框架;
- [science.js](#) - JavaScript中的科学统计计算;
- [Z3d](#) - 在Three.js上轻松地绘制交互式3D图;

- [Sigma.js](#) - JavaScript库，专门用于图形绘制；
- [C3.js](#) - 基于D3.js的定制库，能够轻松绘制图表；
- [ZingChart](#) - Vanilla JS编写的库，用于大数据可视化；
- [cheminfo](#) - 数据可视化和分析的平台，使用[可视化](#)项目。

通用机器学习

- [Convnet.js](#) - ConvNetJS是一个JavaScript库，用于训练深度学习模型[深度学习]；
- [Clusterfck](#) - 为Node.js和浏览器，用Javascript实现的凝聚层次聚类；
- [Clustering.js](#) - Javascript为Node.js和浏览器实现的聚类算法；
- [Decision Trees](#) -使用ID3算法实现NodeJS决策树；
- [figure](#) - K-均值、模糊C均值和凝聚聚类；
- [Node-fann](#) -Node.js绑定的快速人工神经网络库（Fast Artificial Neural Network Library, FANN）；
- [Kmeans.js](#) - K-均值算法用JavaScript的简单实现，用于Node.js和浏览器；
- [LDA.js](#) - Node.js的LDA主题建模；
- [Learning.js](#) - 用Javascript实现逻辑回归/ c4.5决策树；
- [Machine Learning](#) - Node.js的机器学习库；
- [Node-SVM](#) - Node.js的支持向量机；
- [Brain](#) - JavaScript中的神经网络[已弃用]；
- [Bayesian-Bandit](#) - Node和浏览器的贝叶斯实现；
- [Synaptic](#) - 用于Node.js和浏览器的无架构神经网络库；
- [kNear](#) - JavaScript实现的k个最近邻算法，用于监督学习；
- [NeuralN](#) - Node.js的C++神经网络库，其优点是有大量的数据集和多线程训练；
- [kalman](#) - Javascript的卡尔曼滤波器；
- [shaman](#) - Node.js库，同时支持简单和多元的线性回归；
- [ml.js](#) - 用于Node.js和浏览器的机器学习和数值分析工具；
- [Pavlov.js](#) -利用马尔可夫决策过程强化学习；
- [MXNet](#) - 轻量级、便携式、灵活的分布式/深度学习系统，可对动态的、突变数据流调度部署，同时也支持Python、R、Julia、Go、Javascript 等编程语言。

其它

- [sylvester](#) - 用于JavaScript的向量和矩阵数学；
- [simple-statistics](#) - 描述，回归和推断统计的JavaScript实现；用文字实现的JavaScript，没有依赖性，能在所有现代浏览器（包括IE）以及在node.js中工作；
- [regression-js](#) - JavaScript库，包含小二乘法拟合方法的集合，用于寻找数据集的趋势；

- [Lyric](#) - 线性回归库;
- [GreatCircle](#) - 计算大圆距的库。

Julia

通用机器学习

- [MachineLearning](#) - Julia机器学习库;
- [MLBase](#) - 一组支持机器学习算法的发展的函数;
- [PGM](#) - 一个概率图模型的Julia框架;
- [DA](#) - 正则判别分析的Julia组件;
- [Regression](#) - 回归分析算法 (例如线性回归和逻辑回归);
- [Local Regression](#) - 局部回归, 超级流畅!
- [Naive Bayes](#) - Julia朴素贝叶斯方法的简单实现;
- [Mixed Models](#) - 用于装配 (统计) 混合效应模型的Julia组件;
- [Simple MCMC](#) - Julia实现基本的MCMC采样;
- [Distance](#) - Julia距离评估模块;
- [Decision Tree](#) - 决策树分类和回归;
- [Neural](#) - Julia神经网络;
- [MCMC](#) - Julia 的MCMC工具;
- [Mamba](#) - Julia中马尔可夫链蒙特卡罗 (MCMC) 的贝叶斯分析;
- [GLM](#) - Julia的广义线性模型。
- [Online Learning](#)
- [GLMNet](#) - GMLNet的Julia包装版, 适合套索/弹性网模型;
- [Clustering](#) - 集群数据的基本功能: K-均值, DP-均值等;
- [SVM](#) - Julia适用的SVM;
- [Kernal Density](#) - Julia的核密度估计量;
- [Dimensionality Reduction](#) - 降维方法;
- [NMF](#) - 非负矩阵分解的Julia包;
- [ANN](#) - Julia人工神经网络;
- [Mocha](#) - 受Caffe启发, Julia的深度学习框架;
- [XGBoost](#) - Julia中的eXtreme Gradient Boosting 包;
- [ManifoldLearning](#) - 用于流形学习和非线性降维的Julia组件;
- [MXNet](#) - 轻量级、便携式、灵活的分布式/深度学习系统, 可对动态的、突变数据流调度部署, 同时也支持Python、R、Julia、Go、Javascript 等编程语言。
- [Merlin](#) - Julia灵活的深度学习框架。

自然语言处理

- [Topic Models](#) - Julia下的主题模型;

- [Text Analysis](#) - Julia下的文本分析包。

数据分析/数据可视化

- [Graph Layout](#) - 纯Julia实现的图布局算法;
- [Data Frames Meta](#) - DataFrames的元编程工具;
- [Julia Data](#) - 处理表格数据的Julia库;
- [Data Read](#) - 从Stata、SAS、SPSS读取文件;
- [Hypothesis Tests](#) - Julia的假设检验;
- [Gadfly](#) - Julia 灵活的统计制图法;
- [Stats](#) - Julia的统计图检验;
- [RDataSets](#) - 用于装载许多R中可用数据集的Julia包;
- [DataFrames](#) - 处理列表数据的Julia库;
- [Distributions](#) - 概率分布和相关函数的Julia包;
- [Data Arrays](#) - 允许缺失值的数据结构;
- [Time Series](#) - Julia的时间序列工具包;
- [Sampling](#) - Julia的基本采样算法。

其他项/展示

- [DSP](#) - 数字信号处理（滤波，周期图，频谱图，窗函数）;
- [JuliaCon Presentations](#) - JuliaCon的演示文稿;
- [SignalProcessing](#) - Julia写的信号处理工具;
- [Images](#) - Julia的图像库。

Lua

通用机器学习

[Torch7](#)

- [cephes](#) - 由Torch包装的Cephes数学函数库，由Stephen L. Moshier开发，库中提供并包装了从Cephes数学库中的180多个特殊的数学函数，是SciPy的核心，作用广泛;
- [autograd](#) - Autograd自动区分本机Torch代码，受原始的Python版本的启发;
- [graph](#) - Torch的图形包;
- [randomkit](#) - 从Numpy提取的随机数生成包，包装成Torch可用形式;
- [signal](#) - 适用Torch-7的信号处理工具箱，包括FFT、DCT、Hilbert、cepstrums、stft等变换;
- [nn](#) - Torch的神经网络包;
- [nngraph](#) - 此软件包为Torch7 nn库提供图形计算;
- [nnx](#) - 一个不稳定的实验包，能够扩展Torch的内置nn库;

- [rnn](#) - 能够扩展Torch nn库的递推神经网络库，包括RNNs、LSTMs、GRUs、BRNNs、BLSTMs等等；
- [dpnn](#) - 许多有用的功能，并不是主要的神经网络包的一部分；
- [dp](#) - 深度学习库，专门使用Torch7分配进行合理化的研究和开发，它强调通过优雅地使用面向对象的设计模式，从而实现灵活性；
- [optim](#) - 针对Torch的优化库，包括Adagrad、共轭梯度、LBFGS, RProp等算法；
- [unsup](#) - Torch中的无监督学习包，提供了与nn相兼容的模块（LinearPsd、ConvPsd、AutoEncoder等），以及独立的算法（k-均值、PCA）；
- [manifold](#) - 操作流形的包；
- [svm](#) - Torch的支持向量机库；
- [lbfgs](#) - 针对liblbfgs的FFT包；
- [vowpalwabbit](#) - torch的旧式vowpalwabbit接口；
- [OpenGM](#) - OpenGM是一个用于图形化建模和推断的C++库， Lua binding提供了一种简单的描绘图形的方法，用Lua描绘，再用OpenGM优化；
- [sphagetti](#) - @MichaelMathieu开发的用于torch7的Spaghetti（稀疏线型）模块；
- [LuaSHKit](#) - 将局部敏感哈希库SHKit包装成Lua可用形式；
- [kernel_smoothing](#) - KNN，内核加权平均，以及局部线性回归平滑器；
- [cutorch](#) - Torch的CUDA实现；
- [cunn](#) - Torch的CUDA神经网络实现；
- [imgraph](#) - Torch的图像/图形库，该软件包提供了在图像上构造图形、分割、组建树、并转换成图像的例程；
- [videograph](#) - Torch的视频/图形库，该软件包提供了在视频上构造图形、分割、组建树、并转换成视频的例程；
- [saliency](#) - 积分图像的代码和工具，一种基于快速积分直方图找到兴趣点的库；
- [stitch](#) - 使用hugin拼接图像，并将其应用以便视频序列；
- [sfm](#) - 运动场景束调整/结构包；
- [fex](#) - Torch中用于特征提取的软件包，提供SIFT和dSIFT模块；
- [OverFeat](#) - 当前最先进的通用密集特征提取方式。

[Numeric Lua](#)

[Lunatic Python](#)

[SciLua](#)

[Lua - Numerical Algorithms](#)

[Lunum](#)

演示与脚本

[torch7核心演示资源库](#)

- 线性回归，逻辑回归；
- 面部检测器（训练和检测作为独立的演示）；
- 基于mst的分段器；
- train-a-digit-classifier
- train-autoencoder
- optical flow demo
- train-on-housenumbers
- train-on-cifar
- 深层网络追踪；
- Kinect 演示；
- 滤波组件可视化；
- 网络卓越度；

[Training a Convnet for the Galaxy-Zoo Kaggle challenge \(CUDA demo\)](#)

[Music Tagging](#) - Torch7的音乐标记脚本；

[torch-datasets](#) - 用于加载多个大众数据集的脚本，数据集包括如下：

- BSR 500
- CIFAR-10
- COIL
- Street View House Numbers
- MNIST
- NORB

[Atari2600](#) - 从Arcade学习环境的静态帧中生成数据集的脚本。

Matlab

计算机视觉

- [Contourlets](#) - MATLAB源代码，能够实现轮廓变换及其效用函数；
- [Shearlets](#) - 剪切波变换的MATLAB代码；
- [Curvelets](#) - 曲波变换是小波变换的高维推广，用于表示不同尺度和不同角度的图像；
- [Bandlets](#) - bandlet变换的MATLAB代码；
- [mexopencv](#) - 针对OpenCV库，用于收集和开发MATLAB MEX函数的套件；

自然语言处理

- [NLP](#) - 用于Matlab的NLP库；

通用机器学习

- [Training a deep autoencoder or a classifier on MNIST digits](#) - 在MNIST数字上训练深度自编码或深度分类[深度学习];
- [Convolutional-Recursive Deep Learning for 3D Object Classification](#) - 针对三维目标分类的卷积递归深度学习[深度学习];
- [t-Distributed Stochastic Neighbor Embedding](#) - T-分布式随机邻域嵌入 (T-SNE) 法, 已获奖, 该降维技术特别适合高维数据集的可视化;
- [Spider](#) - Matlab中完整的面向对象的环境, 用于机器学习;
- [LibSVM](#) - 支持向量机程序库;
- [LibLinear](#) - 大型线性分类库;
- [Machine Learning Module](#) - 机器学习课程, 包括PDF、讲义及代码;
- [Caffe](#) - 深度学习框架, 开发时考虑到了清洁度, 可读性和速度;
- [Pattern Recognition Toolbox](#) - Matlab中完整的面向对象环境, 用于机器学习;
- [Pattern Recognition and Machine Learning](#) - C. Bishop在书Pattern Recognition and Machine Learning中描述了一个算法, 该软件包包含了书中算法的matlab实现;
- [Optunity](#) - 该库致力于用简单、轻量级的API进行自动超参数优化, 以方便直接替换网格搜索; Optunity用Python编写, 但与MATLAB无缝连接;

数据分析/数据可视化

- [matlab_gbl](#) - 处理图形问题的MATLAB程序包;
- [gamic](#) - 图形算法用纯Matlab高效实现, 以补充MatlabBGL的MEX函数;

.NET

计算机视觉

- [OpenCVDotNet](#) - OpenCV工程封装器, 用于.NET应用程序;
- [Emgu CV](#) - OpenCV的跨平台封装器, 可在Mono中编译, 并运行于Windows、Linux、Mac OS X、iOS、Android等系统;
- [AForge.NET](#) - 开放源码C#框架, 用于计算机视觉和人工智能领域的开发和研究人员, 目前已转移到GitHub发展;
- [Accord.NET](#) - 连同AForge.NET, 可为Windows、Windows RT和Windows Phone提供图像处理 and 计算机视觉算法, 一些组件适用于Java和Android;

自然语言处理

- [Stanford.NLP for .NET](#) - 斯坦福大学NLP包在.NET上的完整移植, 也可作为NuGet包进行预编译;

通用机器学习

- [Accord-Framework](#) - 完整的Accord.NET框架, 用于机器学习, 计算机视觉, 计算机试听, 信号处理和统计应用;

- [Accord.MachineLearning](#) - 该软件包是Accord.NET框架的一部分，包括支持向量机、决策树、朴素贝叶斯模型、K均值、高斯混合模型和一些通用算法，例如适用于机器学习应用的RANSAC、交叉验证和网格搜索；
- [DiffSharp](#) - 一个自动微分（AD）库，能够提供精确高效，用于机器学习和优化应用的导数（梯度、Hessians、雅可比行列式、方向导数、无矩阵Hessian-和雅可比矢量积），操作可以嵌套到任何级别，这意味着你可以通过利用内部微分，来计算精确的高阶导数，如超参数优化函数的应用；
- [Vulpes](#) - F# 写的Deep belief与深度学习的实现，并在Alea.cuBase下用CUDA GPU 执行；
- [Encog](#) - 一种先进的神经网络和机器学习框架。 Encog包含用于创建各类网络的类，同时也支持为神经网络规范和处理数据的类。它的训练采用多线程弹性传播。同时使用GPU 加快处理时间，此外还提供了图形化界面来帮助建模和训练神经网络；
- [Neural Network Designer](#) - DBMS管理系统和神经网络的设计平台，该设计应用程序使用WPF开发，是一个用户界面，允许用户设计自己的神经网络，查询网络，创建和配置能够提出问题、并能从反馈中学习的聊天机器人，聊天机器人甚至可以搜集Internet上的信息，以便在返回时输出，或者用于学习；

数据分析/数据可视化

- [numl](#) - 机器学习库，用于简化预测和聚类的标准建模技术；
- [Math.NET Numerics](#) - Math.NET工程的数值基础，旨在为科学、工程中的数值计算提供了方法和算法，支持 Windows, Linux 和 Mac上的.Net 4.0, .Net 3.5 和 Mono , Silverlight 5, WindowsPhone/SL 8, WindowsPhone 8.1 以及装有 PCL Portable Profiles 47 及 344的Windows 8, 装有 Xamarin的Android/iOS；
- [Sho](#) - 用于数据分析和科学计算的交互环境，可以使用编译代码（.NET）无缝连接脚本（用IronPython），以实现快速、灵活的原型设计；该环境包括线性代数、数据可视化等强大高效的库，可供任何.NET语言使用，还为快速开发提供了功能丰富的交互式shell。

Objective

通用机器学习

- [YCML](#) - 用于Objective-C 和Swift (OS X / iOS)的机器学习框架；
- [MLPNeuralNet](#) - 用于 iOS and Mac OS X的快速多层感知神经网络库，可通过神经网络预测新实例，建立在苹果公司的加速框架之上，如果需要，可利用向量化操作和硬件加速操作；
- [MAMachineLearning](#) - Objective-C多层感知库，通过反向传播全面支持训练，用vDSP和vecLib实现，比同等情况下的Java快20倍，包括从Swift使用的示例代码；
- [BPN-NeuralNetwork](#) - 能够实现3层神经网络（输入层、隐藏层和输出层），其命名来源于BP神经网络（BPN），可用于产品建议、用户行为分析、数据挖掘和数据分析；

- [Multi-Perceptron-NeuralNetwork](#) - 它能够实现基于BPN的多感知神经网络，并设计于无限隐藏层上；
- [KRHebbian-Algorithm](#) - 机器学习神经网络中的非监督和自我学习算法（调整权重）；
- [KRKmeans-Algorithm](#) - 它实现了K-均值聚类 and 分类算法，可用于数据挖掘和图像压缩；
- [KRFuzzyCMeans-Algorithm](#) - 它实现了机器学习中的模糊C均值（FCM）模糊聚类/分类算法，可用于数据挖掘和图像压缩；

OCaml

通用机器学习

- [Oml](#) - 通用的统计学和机器学习库；
- [GPR](#) - OCaml中高效的高斯过程回归算法；
- [Libra-Tk](#) - 用离散概率模型学习和推理的算法。

PHP

自然语言处理

- [jieba-php](#) - 中文分词工具。

通用机器学习

- [PredictionBuilder](#) - 采用线性回归建立预测的机器学习库。

Python

计算机视觉

- [Scikit-Image](#) - Python图像处理算法的集合；
- [SimpleCV](#) - 一个开源的计算机视觉框架，能够访问如OpenCV的高性能计算机视觉库，用Python编写，可运行在Mac、Windows和Ubuntu Linux等操作系统；
- [Vigranumpy](#) - Python绑定的VIGRA C++计算机视觉库；
- [OpenFace](#) - 免费的开源面部识别深层神经网络；

自然语言处理

- [NLTK](#) - 用人类语言数据编写Python程序的一个领先平台；
- [Pattern](#) - Python编程语言的Web挖掘模块，有用于自然语言处理、机器学习等的工具；
- [Quepy](#) - 一个Python框架，将自然语言问题转换成数据库查询语言；
- [TextBlob](#) - 给常规的自然语言处理（NLP）任务提供一致的API，以NLTK和Pattern为基础，并和两者都能很好兼容；
- [YAlign](#) - 语句校准器，较好地用于从比较语料库中提取并行语句；
- [jieba](#) - 中文分词工具；
- [SnowNLP](#) - 用于处理中文文本的库；

- [spammy](#) - 建在NLTK之上的库，用于垃圾邮件过滤；
- [loso](#) - 另一个中文分词库；
- [genius](#) - 基于条件随机域的中文断词库；
- [KoNLPy](#) - 用于韩语自然语言处理的Python包；
- [nut](#) - 自然语言理解工具包；
- [Rosetta](#) - 文本处理工具和包装器（例如Vowpal Wabbit）；
- [BLLIP Parser](#) - Python绑定的BLLIP自然语言解析器（也称为Charniak-Johnson解析器）；
- [PyNLPl](#) - Python的自然语言处理库，Python的通用NLP库，还包含一些特定模块，用于常规的NLP格式分析，尤其是[FoLiA](#)和ARPA语言模型，Moses phrasetales, GIZA++ alignments等；
- [python-ucto](#) - Python绑定的ucto（包括Unicode，基于规则，各种语言的的标记生成器）；
- [python-frog](#) - Python绑定的Frog，用于Dutch的NLP套件（词性标注，词性还原，依存分析，NER）；
- [python-zpar](#) - Python绑定的ZPar，统计性的部分词性标注，constituency解析器和英语依存句法分析；
- [colibri-core](#) - Python绑定的C++库，用于以快速和存储器高效的方式提取并使用基本语言结构，例如-grams 和skipgrams；
- [spaCy](#) - 使用Python和Cython，NLP的产业优势；
- [PyStanfordDependencies](#) - Python接口，用于将Penn Treebank树转换成Stanford Dependencies；

通用机器学习

- [machine learning](#) - 自动构建[网络接口](#)，[编程接口](#)API套件，用于支持向量机；相应的数据集（多个）存储到一个SQL数据库，然后生成用于预测的模型（多个），再被存储到NoSQL数据存储中；
- [XGBoost](#) - Python绑定的 eXtreme Gradient Boosting（树）库；
- [Bayesian Methods for Hackers](#) - 关于Python编程概率规划的图书/
IPythonnotebook；
- [Featureforge](#) - 用于创建和测试机器学习的特点的一组工具，具有兼容scikit学习的API；
- [MLlib in Apache Spark](#) - Spark下的分布式机器学习库；
- [scikit-learn](#) - 基于SciPy的机器学习模块；
- [metric-learn](#) - 测度学习的Python模块；

- [SimpleAI](#) - 实现了许多书*Artificial Intelligence, a Modern Approach*书中描述的人工智能算法，着重于提供一个易于使用、方便存档和测试的库；
- [astroML](#) - 用于天文学的机器学习和数据挖掘；
- [graphlab-create](#) - 实现各种机器学习模型（回归、聚类、推荐系统、图形分析等）的库，基于可以磁盘存储的DataFrame；
- [BigML](#) - 用于连接外部服务器的库；
- [pattern](#) - Python Web挖掘模块；
- [NuPIC](#) - 智能计算的Numenta平台；
- [Pylearn2](#) - 基于[Theano](#)的机器学习库；
- [keras](#) - 基于[Theano](#)的模块化神经网络库；
- [hebel](#) - Python下的GPU加速深度学习库；
- [Chainer](#) - 灵活的神经网络架构；
- [gensim](#) - 主题建模工具；
- [topik](#) - 主题建模工具包；
- [PyBrain](#) - 另一个Python机器学习库；
- [Brainstorm](#) - 继PyBrain后，快速、灵活、有趣的神经网络；
- [Crab](#) - 灵活、快速的推荐引擎；
- [python-recsys](#) - 实现推荐系统的Python库；
- [thinking bayes](#) - 关于贝叶斯分析的书籍；
- [Restricted Boltzmann Machines](#) - Python编写的受限的玻尔兹曼机[深度学习]；
- [Bolt](#) - Bolt在线学习工具箱；
- [CoverTree](#) - cover tree的Python实现，`scipy.spatial.kdtree`便捷的替代；
- [nilearn](#) - Python编写的用于神经影像的机器学习；
- [Shogun](#) - Shogun机器学习工具箱；
- [Pyevolve](#) - 遗传算法框架；
- [Caffe](#) - 深度学习框架，开发时考虑了清洁度、可读性和速度；
- [breze](#) - 基于Theano的用于深层递归神经网络的库；
- [pyhsmm](#) - 用于近似无监督推论的库，存在于贝叶斯隐马尔可夫模型（HMM模型）和隐藏式半马尔可夫模型（HSMMs），HDP-HMM和HDP-HSMM侧重于非参数贝叶斯扩展，大多是弱极限近似值；
- [mr job](#) - 让Python程序运行在Hadoop上的库；
- [SKLL](#) - scikit-learn的包装器，使其实验起来较容易；
- [neurolab](#) - <https://github.com/zueve/neurolab>
- [Spearmint](#) - Spearmint是根据论文中提出的算法执行贝叶斯优化的程序包，论文如下：Practical Bayesian Optimization of Machine Learning Algorithms. Jasper

Snoek, Hugo Larochelle and Ryan P. Adams. Advances in Neural Information Processing Systems, 2012.

- [Peb1](#) - 贝叶斯学习的Python环境;
- [Theano](#) - 在Python中, 用优化的GPU元编程代码生成阵列导向的优化数学编译器;
- [TensorFlow](#) - 使用数据流图进行数值计算的开源软件库;
- [yahmm](#) - Python下的隐马尔可夫模型, 考虑到速度和效率, 用Cython来实现;
- [python-timbl](#) - 一个Python扩展模块, 可包装完整的TiMBL C++编程接口; Timbl是一个详尽的k最近邻域机器学习工具包;
- [deap](#) - 进化算法框架;
- [pydeep](#) - Python下的深度学习;
- [mlxtend](#) - 工具库, 包括对数据科学和机器学习任务等非常有用的工具;
- [neon](#) - Nervana基于Python的[高性能](#)深度学习框架[深度学习];
- [Optunity](#) - 该库致力于用一个简单、轻量级的API直接替换网格搜索以实现自动超参数优化;
- [Neural Networks and Deep Learning](#) - 《神经网络和深度学习》一书中的代码示例[深度学习];
- [Annoy](#) - 近似最近邻算法的实现;
- [skflow](#) - 为TensorFlow简化的界面, 模仿了Scikit Learn;
- [TPOT](#) - 使用遗传编程自动创建并优化了机器学习流水线的工具;
- [pgmpy](#) - 用于概率图解模型的Python库;
- [DIGITS](#) - 深度学习GPU训练系统 (DIGITS) 是用于训练深度学习模型的Web应用程序;
- [Orange](#) - 同时适用于初学者和专家的开源数据可视化和数据分析;
- [MXNet](#) - 轻量级、便携式、灵活的分布式/深度学习系统, 可对动态的、突变数据流调度部署, 同时也支持Python、R、Julia、Go、Javascript 等编程语言。
- [milk](#) - 强调监督分类法的机器学习工具包;
- [TFlearn](#) - 针对TensorFlow的高层次API深度学习库。

数据分析/数据可视化

- [SciPy](#) - 基于Python的数学、科学、工程开源软件生态系统;
- [NumPy](#) - Python科学计算基础包;
- [Numba](#) - Python的低级虚拟机JIT编译器, 由Cython and NumPy的开发者编写, 供科学计算使用;
- [NetworkX](#) - 用于复杂网络使用的高效率软件;
- [igraph](#) - 绑定igraph 库-通用图形库 ;
- [Pandas](#) - 该库提供了高性能、易用的数据结构及数据分析工具;
- [Open Mining](#) - Python中的商业智能工具 (Pandas web接口) ;

- [PyMC](#) - Markov Chain Monte Carlo (MCMC) 采样工具包;
- [zipline](#) - Python的算法交易库;
- [PyDy](#) - 全名Python Dynamics, 协助基于NumPy, SciPy, IPython以及 matplotlib 的动态建模 workflow;
- [SymPy](#) - Python中的符号数学库;
- [statsmodels](#) - Python中的统计建模及计量经济学库;
- [astropy](#) - Python中的天文学程序库;
- [matplotlib](#) - Python中的2D绘图库;
- [bokeh](#) - Python的交互式Web绘图库;
- [plotly](#) - Python 和matplotlib的协作web绘图库;
- [vincent](#) - Python到 Vega的转换器;
- [d3py](#) - Python的绘图库, 基于[D3.js](#);
- [ggplot](#) - 和R语言里的ggplot2提供同样的API;
- [ggfortify](#) - 统一到ggplot2流行的R包的接口;
- [Kartograph.py](#) - Python中渲染SVG图的库, 效果出众R;
- [pygal](#) - Python的SVG图表生成器;
- [PyQtGraph](#) - 基于PyQt4/PySide和NumPy 纯Python编写的图形和GUI库;
- [pycascading](#)
- [Petrel](#) - 纯Python中用于编写、提交、调试和监视Storm拓扑结构的工具;
- [Blaze](#) - NumPy和Pandas的大数据接口;
- [emcee](#) - Python集成采样工具包用于仿射不变的MCMC;
- [windML](#) - 用于风能分析与预测的Python框架;
- [vispy](#) - 基于GPU的高性能交互式OpenGL 2D/3D数据可视化库;
- [cerebro2](#) - NuPIC的一个基于Web的可视化和调试平台;
- [NuPIC Studio](#) - 一个all-in-one NuPIC 分层时间可视化和调试的超级工具!
- [SparklingPandas](#) Pandas on PySpark (POPS)
- [Seaborn](#) - 基于matplotlib的Python可视化库;
- [bqplot](#) - Jupyter (IPython)中用于绘图的API;
- [pastalog](#) - 用于训练神经网络的简单、实时可视化的服务器;
- [caravel](#) - 可视、直观、互动的数据探索平台。

杂项脚本/iPython笔记/代码库

- [BioPy](#) - Python中的生物启发和机器学习算法;
- [pattern_classification](#)
- [thinking_stats_2](#)
- [hyperopt](#)

- [numpic](#)
- [2012-paper-diginorm](#)
- [A gallery of interesting IPython notebooks](#)
- [ipython-notebooks](#)
- [data-science-ipython-notebooks](#) - 持续更新的数据科学Python笔记: Spark、Hadoop MapReduce、HDFS、AWS、Kaggle、scikit-learn、matplotlib、pandas、NumPy、SciPy、和其他各种命令行;
- [decision-weights](#)
- [Sarah Palin LDA](#) - Sarah Palin关于主题建模的电子邮件;
- [Diffusion Segmentation](#) - 基于扩散方法的图像分割算法集合;
- [Scipy Tutorials](#) - SciPy教程, 该部分已过时, 请查看scipy-lecture-notes;
- [Crab](#) - Python的推荐引擎库;
- [BayesPy](#) - Python中的贝叶斯推断工具;
- [scikit-learn tutorials](#) - scikit-learn学习笔记本系列;
- [sentiment-analyzer](#) - 推特情绪分析器;
- [sentiment_classifier](#) - 采用词义消歧的情感分类器;
- [group-lasso](#) - 一些坐标下降算法实验, 应用于(稀疏)群套索模型;
- [jProcessing](#) - Kanji / Hiragana / Katakana 到Romaji的转换器。Edict 字典和平行句搜索。两个JP句间相似度。日文情感分析。在Python运行 Cabocha (ISO 8859-1配置);
- [mne-python-notebooks](#) -使用 mne-python进行EEG/MEG数据处理的IPython笔记;
- [pandas cookbook](#) - 使用Python pandas库的方法书;
- [climin](#) - 机器学习的优化程序库, 用Python实现了梯度下降、LBFGS、rmsprop、adadelat 等算法;
- [Allen Downey's Data Science Course](#) - 2014春季Olin挑战数据科学源码;
- [Allen Downey's Think Bayes Code](#) - Think Bayes 的代码库;
- [Allen Downey's Think Complexity Code](#) - Allen Downey's book Think Complexity 源码;
- [Allen Downey's Think OS Code](#) - Think OS的文本和对应源码: 操作系统简介;
- [Python Programming for the Humanities](#) - 为那些对应知识人开设的Python编程课。重点在文本处理/NLP;
- [GreatCircle](#) - 计算大圆距离的库;
- [Optunity examples](#) - 如何将Optunity 和机器学习库协同使用的演示示例;
- [Dive into Machine Learning with Python Jupyter notebook and scikit-learn](#) - “我最初最为黑客来学习Python, 越陷越深。我想利用在机器学习中使用它。如

何您对此感兴趣，请加入我们。”

- [TDB](#) - TensorDebugger (TDB) 是深度学习可视化调试器，使用断点调试、互动和可视化扩展 [TensorFlow](#)。

Ruby

自然语言处理

- [Treat](#) - 文本检索与注释工具包，目前为止我见过Ruby上的最全面的工具包；
- [Ruby Linguistics](#) - Linguistics框架可以用任何语言为Ruby对象构建语言学工具，它包括一个语言无关的通用前端、一个将语言代码映射到语言名的模块以及一个含有多英文语言工具的模块；
- [Stemmer](#) - 使Ruby可以用 libstemmer_c中的接口；
- [Ruby Wordnet](#) - WordNet的Ruby接口库；
- [Raspell](#) - raspell是一个绑定到Ruby的接口；
- [UEA Stemmer](#) -UEALite Stemmer的Ruby移植版，用来供搜索和检索用的保守的词干分析器 ；
- [Twitter-text-rb](#) - 一个可以将推特中的用户名、列表和话题标签自动连接并提取出来的库。

通用机器学习

- [Ruby Machine Learning](#) - Ruby语言实现的一些机器学习算法；
- [Machine Learning Ruby](#)
- [jRuby Mahout](#) - JRuby Mahout一个精华！在JRuby世界中释放了Apache Mahout的威力；
- [CardMagic-Classifier](#) - 可用贝叶斯及其他分类法的通用分类器模块；
- [rb-libsvm](#) - LIBSVM的Ruby语言绑定，是一个支持向量机的库。

数据分析/数据可视化

- [rsruby](#) - Ruby - R的桥梁；
- [data-visualization-ruby](#) - 关于数据可视化的Ruby Manor演示的源代码和支持内容；
- [ruby-plot](#) - 将gnuplot封装为Ruby形式，尤其适合将ROC曲线转化为svg文件；
- [plot-rb](#) - 基于Vega和D3的ruby绘图库；
- [scruffv](#) - Ruby中的出色的图形工具包；
- [SciRuby](#)
- [Glean](#) - 数据管理工具；
- [Bioruby](#)
- [Arel](#)

杂项

- [Big Data For Chimps](#)
- [Listof](#) - 基于数据收集的社区，在创业板块。获取更多资料的TXT、Json或Hash列表。 [Demo/Search for a list](#)

Rust

通用机器学习

- [deeplearn-rs](#) - deeplearn-rs 提供了使用矩阵乘法、加法简单的网络，遵循MIT协议；
- [rustlearn](#) - 支持逻辑回归、支持向量机、决策树和随机森林的机器学习框架；
- [rusty-machine](#) - 一个纯rust语言机器学习库；
- [leaf](#) - 用于机器智能的开源框架，其灵感来自TensorFlow和 Caffe，遵循 MIT 协议；
- [RustNN](#) - RustNN 是一个前向反馈神经网络库。

R

通用机器学习

- [ahaz](#) - ahaz: 用于正则化半参数风险回归模型；
- [arules](#) - arules:[频繁项集与关联规则的挖掘](#)；
- [bigrf](#) - bigrf:大随机森林：大型数据集的分类和回归分析；
- [bigRR](#) - bigRR: 广义岭回归 （在 $P \gg N$ 的情况特具有特殊优势）；
- [bmrn](#) - bmrn: 正则化风险最小方案集；
- [Boruta](#) - Boruta: 一种用于所有相关特征选择的包装算法；
- [bst](#) - bst: 梯度推进（Gradient Boosting）；
- [C50](#) - C50: C5.0决策树和基于规则模型；
- [caret](#) - 分类与回归训练:R语言150个机器学习算法的统一接口；
- [caretEnsemble](#) - 用于拟合和创建多符号模型的框架 ；
- [Clever Algorithms For Machine Learning](#)（机器学习的巧妙算法）
- [CORElearn](#) - 分类、回归、特征评价与序评价；
- [CoxBoost](#) - CoxBoost: 基于提升单一存活点或竞争风险Cox模型；
- [Cubist](#) - Cubist: 基于规则和实例的回归模型；
- [earth](#) - 多元自适应回归样条模型；
- [evtree](#) - evtree: 全局最优树的进化学习；
- [fpc](#) - 灵活的聚类程序；
- [frbs](#) - frbs: 基于模糊规则的分类和回归任务系统；

- [GAMBoost](#) - 基于似然法的广义线性模型和加性模型;
- [gamboostLSS](#) - GAMLSS的Boosting方法;
- [gbm](#) - 广义回归模型;
- [glmnet](#) - 用于实现套索和弹性网络化的广义线性模型
- [glmnet](#) - Cox比例风险模型和广义线性模型的L1正则化路径;
- [GMMBoost](#) - 广义混合模型中基于似然估计的Boosting;
- [grplasso](#) - 采用群体套索惩罚的用户指定模型;
- [grpreg](#) - 分组协变量回归模型的正则化路径;
- [h2o](#) - 大规模快速、并行、分布式机器学习算法框架—深度学习、随机森林、GBM、KMeans、PCA、GLM;
- [hda](#) - 异方差判别分析;
- 统计学习概论 ([Introduction to Statistical Learning](#))
- [ipred](#) - ipred: 改进的预测模型;
- [kernlab](#) - kernlab: 基于内核的机器学习实验室;
- [klaR](#) - klaR: 分类和可视化;
- [lars](#) - lars: 最小角回归、套索和向前逐步回归;
- [lasso2](#) - lasso2: L1约束估计又名“套索”;
- [LiblinearR](#) - LiblinearR: 基于LiblinearC / C++ 库线性预测模型;
- [LogicReg](#) - LogicReg: 逻辑回归;
- [Machine Learning For Hackers](#)
- [maptree](#) - 映射、剪切和图形树模型;
- [mboost](#) - 基于模型的Boosting
- [medley](#) - 使用贪婪逐步算法的混合回归模型;
- [mlr](#) - R语言中的机器学习;
- [mvpart](#) - 多元分割
- [ncvreg](#) - SCAD的正则化路径和MCP惩罚回归模型;
- [nnet](#) - 前向反馈神经网络和多项式对数线性模型;
- [oblique.tree](#) - 用于数据分类的Oblique Trees ;
- [pamr](#) - 微阵列预测分析;
- [party](#) - 一个递归Partytioning的实验室;
- [partykit](#) - 一个递归Partytioning的工具包;
- [penalized](#) - 在广义线性模型和Cox模型中L1 (套索和融合套索) 和L2 (岭) 惩罚估计;
- [penalizedLDA](#) - 采用线性判别进行惩罚分类;
- [penalizedSVM](#) - 采用惩罚功能的特征选择向量机;

- [quantregForest](#) - 分位数回归森林算法;
 - [randomForest](#) - 用于分类和回归的Breiman和Cutler随机森林算法;
 - [randomForestSRC](#) - 用于生存、回归和分类的随机森林算法 (RF-SRC);
 - [rattle](#) - R中的数据挖掘的图形用户界面;
 - [rdetools](#) - 用于在特征空间的相关维度估计;
 - [REEMtree](#) - 纵向 (面板) 数据的随机效应回归树;
 - [relaxo](#) - 放松的套索;
 - [rgenoud](#) - R语言版的遗传优化程序包;
 - [rgp](#) - R 语言中的遗传程序框架;
 - [Rmalschains](#) - R语言中基于局部搜索链的模因算法进行持续优化算法;
 - [rminer](#) - rminer: 在分类和回归中数据挖掘 (例如NN、SVM) 方法的简单应用;
 - [ROCR](#) - 用于评分分类器性能的可视化;
 - [RoughSets](#) - 采用粗糙集与模糊粗糙集理论的数据分析包;
 - [rpart](#) - 递归分割与回归树;
 - [RPMM](#) - 递归分割混合模型;
 - [RSNNS](#) - R中的神经网络, 采用了Stuttgart神经网络模拟器;
 - [RWeka](#) - R/Weka 接口;
 - [RXshrink](#) - 广义岭或最小角回归的最大似然收缩;
 - [sda](#) - 收缩判别分析和CAT得分变量选择;
 - [SDDA](#) - 逐步对角判别分析;
 - [SuperLearner](#) and [subsemble](#) - 多算法集成学习包;
 - [svmpath](#) - SVM向量路径算法;
 - [tgp](#) - 高斯型贝叶斯模型;
 - [tree](#) - 分类与回归树;
 - [varSelRF](#) - varSelRF: 使用随机森林的变量选择;
 - [XGBoost.R](#) - 结合R的极值梯度Boosting 库;
 - [Optunity](#) - 通过一个简单、轻量的API可方便更换网格搜索, 用于自动化参数优化的库, [Optunity](#) 是由Python编写而成但无缝衔接到R;
 - [igraph](#) - 通用图形库;
 - [MXNet](#) - 轻量级、便携式、灵活的分布式/深度学习系统, 可对动态的、突变数据流调度部署, 同时也支持Python、R、Julia、Go、Javascript 等编程语言。
- 数据分析/数据可视化
- [ggplot2](#) - 一个基于图形语法的数据可视化包。

SAS

通用机器学习

- [Enterprise Miner](#) - 通过使用GUI或代码创建可部署模型进行数据挖掘和机器学习;
- [Factory Miner](#) - 通过使用GUI, 横跨多个市场或客户自动创建可部署机器学习模型。

数据分析/数据可视化

- [SAS/STAT](#) - 用于高级统计分析;
- [University Edition](#) - 免费! 其中包括数据分析和可视化的必要SAS软件包, 以及在线SAS课程。

高性能机器学习

- [High Performance Data Mining](#) - 通过使用GUI或代码在MPP环境(包括Hadoop)创建可部署模型进行数据挖掘和机器学习;
- [High Performance Text Mining](#) - 通过使用GUI或代码在MPP环境(包括Hadoop)进行文本挖掘。

自然语言处理

- [Contextual Analysis](#) - 通过GUI向非结构化文本增加结构;
- [Sentiment Analysis](#) - 通过GUI从文本抽取情感信息;
- [Text Miner](#) - 通过GUI或代码进行文本挖掘。

示例和脚本

- [ML Tables](#) - 包含机器学习最佳实践的简明速查表;
- [enlighten-apply](#) - SAS机器学习技术应用的相关代码和资料;
- [enlighten-integration](#) - 用于说明SAS与Java、PMML、Python和R中的分析技术相结合的代码和材料;
- [enlighten-deep](#) - 用于说明SAS中神经网络及其多个隐藏层的相关代码和材料;
- [dm-flow](#) - 关于SAS Enterprise Mine处理流程图的库, 通过具体的数据挖掘实例帮助您学习。

Scala

自然语言处理

- [ScalaNLP](#) - ScalaNLP一系列机器学习和数值计算库的集合;
- [Breeze](#) - Breeze是Scala中的数值处理库;
- [Chalk](#) - Chalk是一个自然语言处理库。
- [FACTORIE](#) - FACTORIE是可部署的概率建模工具包, 用Scala实现的软件库。为用户提供简洁的语言来创建关系因素图, 评估参数并进行推断。

数据分析/数据可视化

- [MLlib in Apache Spark](#) - Spark中的分布式机器学习库；
- [Scalding](#) - Cascading的Scala AP接口；
- [Summing Bird](#) - 使用Scalding 和 Storm进行Streaming MapReduce；
- [Algebird](#) - Scala中的抽象代数工具；
- [xerial](#) - Scala中的数据管理工具；
- [simmer](#) - 一个进行代数聚合的unix过滤器，用于简化你的数据；
- [PredictionIO](#) - PredictionIO是一个供软件开发者和数据工程师用的机器学习服务器；
- [BIDMat](#) - 用于支持大规模探索性数据分析的CPU和GPU加速矩阵库；
- [Wolfe](#) - 描述性机器学习（Declarative Machine Learning）；
- [Flink](#) - 分布式数据流处理和批量数据处理的开源计算平台；
- [Spark Notebook](#) - 使用Scala和Spark的互动数据科学。

通用机器学习

- [Conjecture](#) - Scalding中的可扩展的机器学习框架；
- [brushfire](#) - Scala中的分布式决策树集成学习工具；
- [ganitha](#) - 基于scalding的机器学习程序库；
- [adam](#) - 一个使用Apache Avro、Apache Spark 和 Parquet的基因组处理引擎，有其专用的文件格式，采用Apache 2.0协议许可；
- [bioscala](#) - Scala编程语言可用的生物信息库；
- [BIDMach](#) - 用于CPU和GPU加速的机器学习库；
- [Figaro](#) - 用于构造概率模型的Scala库；
- [H2O Sparkling Water](#) - 用于H2O 和Spark 交互操作；
- [FlinkML in Apache Flink](#) - Flink中的分布式机器学习库；
- [DynaML](#) - 用于机器学习的Scala库/REPL。

Swift

通用机器学习

- [Swift AI](#) - Swift语言编写的高度优化的人工智能和机器学习；
- [BrainCore](#) - 支持iOS和 OS X的神经网络框架；
- [DeepLearningKit](#) - 支持苹果iOS OS X和tvOS系统的开源深度学习框架。目前支持在苹果操作系统上利用Caffe训练深度卷积网络模型；
- [AIToolbox](#) - A toolbox framework of AI modules written in Swift: Swift语言编写的人工智能（AI）模型工具框架：图/树、线性回归、支持向量机、神经网络、PCA、KMeans、遗传算法、MDP和 混合高斯模型。

