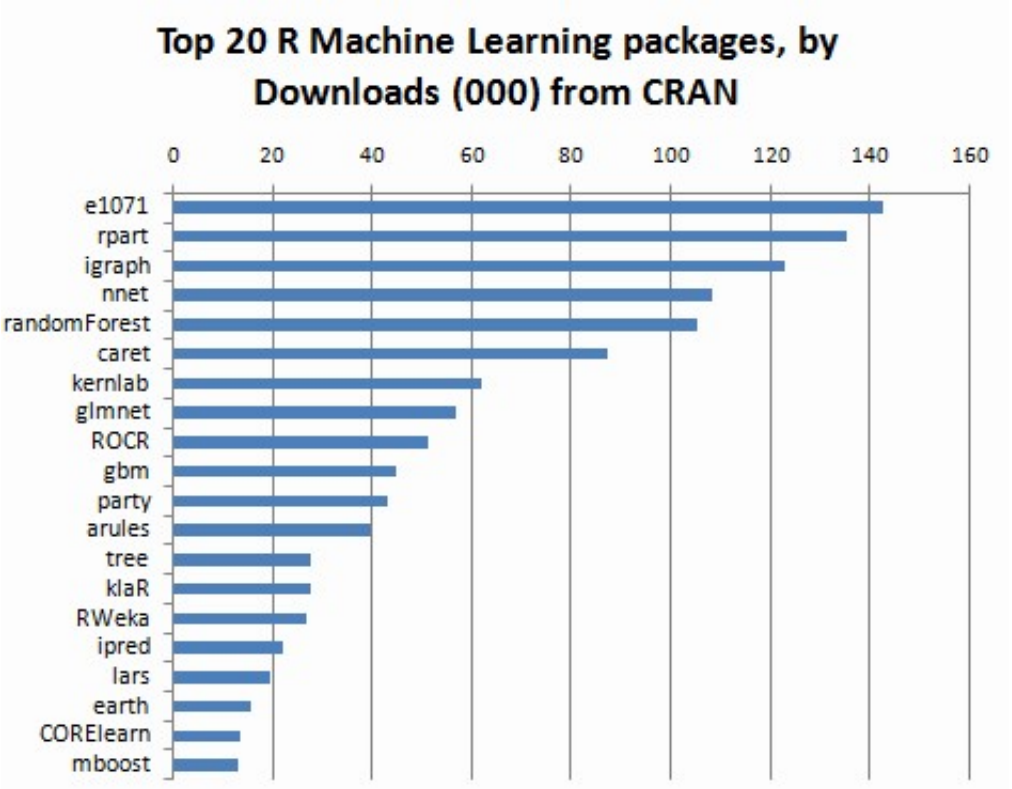
20个在机器学习和数据科学中最常用的R语言包

我们通过分析从2015年1月至5月下载次数最多的R包，列出了前20名流行的机器学习R包。

大多数R包都深受Kagglers大神的最爱，也被资深的笔者所赞美，而这些包的使用率或评价高低不仅仅取决于其它的包对于这个

这个包的依赖程度。还也取决于Crantastic.org并使用其众包能解决方案的用户。但是，用户评价太低以至于不能基于分析来做决定。

我们根据从一月到五月的下载量，通过分析CRAN的日均下载量来统计有多少关于机器学习的包被下载了。



1、e1071

潜类分析函数，短暂性的傅里叶变化、模糊的集群、支持向量机、最短路径计算、装袋集群、朴素贝叶斯分类器等（下载量：142470）。

2、rpart

递归分割和回归树（下载量：135390）。

3、igraph

网络分析工具的集合（下载量：122930）。

4、nnet

前馈神经网络和多元对数线性模型（下载量：108298）。

5、randomForest

Breiman and Cutler的关于分类和回归的随机森林（下载量：105375）。

6、caret

一套试图简化创建预测模型的函数集（下载量：87151）。

7、kernlab

基于内核的机器学习实验室（下载量：62064）。

8、glmnet

Lasso和弹性网正规化广义线性模型（下载量：56948）。

9、ROCR

可视评分分类的操作（下载量：51323）。

10、gbm

产生改良的回归模型（下载量：44670）。

11、party

一个递归分割实验室（下载量：43290）。

12、arules

关联规则挖掘和频繁项集（下载量：39654）。

13、tree

分类和回归树（下载量：27882）。

14、klaR

分类和可视化操作（下载量：27828）。

15、RWeka

关于R或Weka的接口（下载量：26973）。

16、ipred

改良的预测变量（下载量：22358）

17、lars

最小回归角、Lasso和逐步回归分析阶段（下载量：19691）。

18、earth

多元自适应回归样条分析。（下载量：15901）。

19、COREIearn

分类、回归、功能评估和有序评估（下载量：13856）。

20、mboost

基于模型的Booting（下载量：13078）。

你会发现记住这些越来越受人追捧的开源R工具如Rattle，R的一种图形用户界面的数据挖掘工具（下载量：35539），以及fastcluster，R和Python（下载量：14214）的快速的分级聚类程序。

英文原文：Top 20 R Machine Learning and Data Science packages

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（过往记忆）所有，未经许可不得转载。
本文链接:【】（）