

微软打算带着R语言搞仕途？

作者：编译 | zyy 编辑：覃里 2017-01-13 10:00 IT168网站 原创  分享

【IT168 评论】自从微软喊出开源口号的那一天，就一改往日的傲娇，除了自己参与开源还收购了一些开源公司，最有趣的当属收购Revolution Analytics，Revolution Analytics是R语言的开发公司，Revolution Analytics能够提供企业级基于R语言的数据解决方案，可以覆盖大量的数据库和Hadoop系统，同时Revolution Analytics也持续在ParallelR, RHadoop等方面支持R语言开源社区。



被微软收购之后，重新上路的R Server已经成为内部部署和云数据之间的桥梁。微软公司还宣布了一系列R语言工具的更新。R语言已经成为其数据战略的重要组成部分，在Azure和SQL Server中均提供了支持(这就是要搞仕途啊!)。更重要的是，在Azure机器学习服务中，在数据提交到机器学习管道之前，R语言用于预处理数据。可见，R俨然成为微软关键的跨平台产品之一。

R在微软的生态系统中无处不在

开源的R已经成为数据科学的关键工具(根据IEEE，它目前在所有语言中排名第五)，这其中统计学专家对R格外青睐有加，因为Comprehensive R Archive Network(CRAN，一个R语言应用程序的公共库)现在有超过9,000个统计模块和算法可以使用。

微软对R的愿景是打破台式机，本地服务器和云之间的界限。在本地，有一个免费的R语言客户端，以及微软(付费)旗舰版Visual Studio开发环境中的R支持。在内部，R Server可以在Windows和Linux上同时运行。在SQL Server内部，可以访问统计分析工具以及数据，同时支持基于Hadoop和Spark的本地大数据服务。在Azure上，可以和Microsoft的HDInsight服务一起运行R Server。

R是数据科学家的工具。虽然R语言相对简单，但需要对统计分析有深入的了解才能从中获得最大利益。因为许多基本概念需要对复杂统计函数进行研究层面的理解。问题不在于能否编写R代码，而在于它是否能理解你的思想。

所有使用大数据的企业面临的最大问题是：如何获取所需的技能并生成正确的分析结果。更重要的是，可以对分析结果有一个直观的解释，R语言内置的图形工具是可视化分析的关键工具。

Microsoft R Server出道

Microsoft R Server是一款基于R的企业级大数据分析平台，支持各种大数据统计分析，预测性建模和机器学习功能。它是一个十分有用的工具，可以快速尝试新的分析算法和探索使用数据回答问题。这种方法适合作为整个分析生命周期的一部分，从数据准备开始，然后是模型开发，最后将模型转换为可构建到业务应用程序中的工具。

机器学习同样有R

R也是基于GPU的机器学习工具。R可用于帮助训练模型，Microsoft将自己的机器学习算法与最新的R Server版本捆绑在一起，因此你可以在模型上传到本地大数据实例或云之前进行测试。在最近的一次新闻发布会上，Microsoft在天文图像上演示了这种方法，在具有星系库的本地服务器上训练基于机器学习的分类器，然后在云托管的GPU上运行结果模型。



R是一种非常便携的语言，可设计用于处理离散的数据样本。这使它可扩展和解决数据并行问题。相同的R模型可以在多个服务器上运行，可以快速处理大量数据。你需要做的就是适当分割数据，然后将其传递到各种R服务器实例。类似地，相同的代码可以以不同的实现方式运行，因此针对本地数据源构建和测试的模型可以部署在SQL Server数据库内并针对Hadoop数据湖运行。

R使数据模型操作更容易

R非常容易实现，数据科学团队构建需要的模型，开发人员编写应用程序并构建可利用其代码的基础架构就可以了。一旦准备就绪，该模型可以快速部署，应用过程中也可以及时更换模型，而不影响应用程序的其余部分。总之，R允许你以相同的方式，相同的模型在不同的应用中使用相同的数据。

通用模型可以对数据主动响应，例如，当模型预测到航班延迟时，可向航空公司乘客提供延迟和重新预订信息。该模型在获得更多数据之后可以进行细化，从而降低风险。

将R支持构建到SQL Server中有很大意义。随着Microsoft的数据库平台成为内部部署和云数据以及大数据工具系统之间的桥梁，在数据库中使用细粒度的分析工具是一个明智的决定。数据库开发人员可以与数据分析团队一起实施这些模型，而不需要学习任何新技能来将其构建到应用程序中。

微软意识到，不是每个企业都需要或有预算来雇佣数据科学家。如果你正在处理常见的分析问题，例如尝试预测客户流失或检测在线商店中的欺诈行为，你可以选择包含即用模型的SQL Server R服务中的一系列预定义模板。从Microsoft的MSDN可以获得，它们可以在任何R兼容的IDE中自定义，并且可以使用PowerShell脚本部署。

 0

相关文章

[盛拓传媒简介](#) [关于IT168](#) [广告服务](#) [使用条款](#) [投稿指南](#) [联系](#)

北京盛拓优讯信息技术有限公司. 版权所有 中华人民共和国增值电信业务经营许可证 编号：京B2-20170206 北京市

广播电视节目制作经营许可证:(京)字第07177号 信息系统安全等级保护备案：11010813655-00001 测绘资质证书:乙测资

违法和不良信息举报电话 :010-59548436,010-59544810,17352615267,17816876620,niuxiaotong@