

# Revolution R Enterprise 7.4 README

Revolution R Enterprise 7.4 for 64-bit Windows and Linux (Red Hat Enterprise Linux 5.x and 6.x and SUSE Linux Enterprise Server 11.x) is the fastest, most cost effective, enterprise-class, big data big analytics software available for workstations, servers, and distributed platforms such as Hadoop. Supporting a variety of big data statistics, predictive modeling, and machine learning capabilities, Revolution R Enterprise is also 100% R. The 7.4.0 release extends the functionality of R 3.1.3 and includes updated releases of the RevoScaleR package for scalable, fast (multicore), and extensible data analysis with R and the RevoPemaR package for developing your own parallel external memory algorithms in R.

Download and installation instructions are provided in your welcome e-mail.

## What's New in Revolution R Enterprise 7.4

### RevoR

#### *Installer Enhancements*

- Revolution R Enterprise 7.4 can now be installed into an existing R 3.1.3 installation after installing Revolution R Connector 7.4.

#### **New and Removed Packages**

- The Revobase and RevobaseEnt packages have been removed.
- The new RevoUtils package contains basic utility functions such as 'readNews'.
- The new RevoUtilsMath package contains functions to get and set the MKL threads.
- The new RevoIOQ package contains the RevoIOQ function for running an installation and operational qualification test suite on a single machine.

### RevoScaleR

#### *Two-process Architecture*

- RevoScaleR version 7.4 now features a two-process architecture—R itself runs in one process, while RevoScaleR's underlying computational engine runs in a separate process. For RRE 7.4, this is primarily an infrastructure investment, but we expect to take advantage of the new architecture in future releases by providing exciting new interfaces to the computational engine.

#### *Naïve Bayes Classifier*

- The new function rxNaiveBayes performs classification using Bayes Theorem to determine the probability that an observation belongs to a certain class

#### *Stepwise Fit Improvements*

- A new argument, keepStepCoefs, has been added to the rxStepControl function. If TRUE, a data frame stepCoefs will be returned with the fitted model with rows

corresponding to the coefficients and columns corresponding to the iterations. Additional computation may be required to generate the coefficients at each step. Those stepwise coefficients can be visualized by plotting the fitted model with rxStepPlot.

#### ***RxHadoopMR Compute Context***

- The RxHadoopMR compute context now supports additional distributions of Hadoop, including Cloudera CDH 5.2 and 5.3, Hortonworks HDP 2.2, and MapR 4.0.2.

#### ***Efficiency Improvements***

- Import of wide data is now faster; up to 180 times faster on sample data sets with 15 thousand variables.
- Fitting of tree-based models is now faster in most cases because splits are found in parallel. (In the relatively rare case where you are building shallow trees on data sets with few variables, it can be faster to set the findSplitsInParallel flag to FALSE.)
- Prediction on rxDForest and rxBTrees objects has been speeded up by performing the prediction in a single pass over the data.

#### ***Deprecated and Obsolete Functionality***

- The rxGetVarInfoXdf and rxGetInfoXdf functions in RevoScaleR are now deprecated. Use rxGetVarInfo and rxGetInfo.
- Windows XP and Windows Server 2003 are no longer supported.

#### ***Known Issues***

- [\*Known Issues in Revolution R Enterprise 7.4\*](#)