Predictive Analytics with TIBCO Spotfire and TIBCO Enterprise Runtime for R
# Table of Contents

<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Predictive Analytics with TIBCO Spotfire</td>
</tr>
<tr>
<td>4</td>
<td>TIBCO Spotfire Statistics Services</td>
</tr>
<tr>
<td>8</td>
<td>TIBCO Enterprise Runtime for R (TERR)</td>
</tr>
<tr>
<td>11</td>
<td>Predictive Modeling Tools</td>
</tr>
<tr>
<td>13</td>
<td>For more information</td>
</tr>
</tbody>
</table>
Predictive Analytics with TIBCO Spotfire is the premier data discovery and analytics platform, which provides powerful capabilities for our customers, such as dimension-free data exploration through interactive visualizations, and data mashup to quickly combine disparate data to gain insights masked by data silos or aggregations.

Another key strength of TIBCO Spotfire platform is broad predictive analytic functionality. Predictive analytics have entered the mainstream of business analytics in the last few years, and generally can be described as improving future decision-making by enabling learning from an organization’s past collective experience.

Benefits of TIBCO Spotfire and Predictive Analytics

• Easily provide targeted, relevant predictive analytics to business users
  ○ Ensure compliance and proper usage
  ○ Get the answer when needed
• Increase confidence and effectiveness in decision-making
  ○ Reduce uncertainty
  ○ Discover meaningful patterns, important data
  ○ Maximize ROI
• Anticipate and react to emerging trends
• Reduce/manage risk
  ○ Use scenario planning, forecasts, and fraud detection
• Forecast specific behavior, preemptively act on it
  ○ Increase upsell, decrease churn

Predictive Analytics in the Spotfire Platform

There are three main aspects of predictive analytics in the Spotfire platform:

• TIBCO Spotfire Statistics Services (TSSS) provides a predictive analytics ecosystem and enables seamless integration of your existing investments in R, S+, SAS, MATLAB into Spotfire and custom applications, as well as leveraging in-database predictive analytics through Teradata Aster to empower more effective decision-making across your organization
• TIBCO Enterprise Runtime for R (TERR) provides an enterprise-class environment for running R scripts and packages, both within Spotfire and across an organization, enabling you to combine the agility of open source R with the speed and reliability of an enterprise platform
• Predictive Modeling Tools in Spotfire provide deep predictive insights into your data as part of ad hoc analysis without requiring any statistical programming
TIBCO Spotfire Statistics Services

With TIBCO Spotfire Statistics Services, technical and business professionals gain the benefits of a full predictive analytics ecosystem. They gain more confidence in their decisions by using the latest, most relevant predictive analytics available in R, S+, SAS, or MATLAB – without requiring deep expertise in statistics. Organizations increase efficiency by leveraging their existing investments in predictive analytics, giving decision makers self-service access to easy-to-interpret analytic results through Spotfire applications. Scarce statistical resources deploy and control access to a centralized repository of R, S+, SAS, or MATLAB functions, ensuring only the most appropriate and statistically valid analytic methods are used.

For advanced users, TIBCO Spotfire Statistics Services also complements and enhances the usage of S+ and R by allowing statisticians to easily visualize the results of their models and analysis and to deploy these models inside Spotfire applications from a central location.

Predictive Analytics Ecosystem
Benefits

• Easily provide targeted, relevant advanced analytics to large, diverse communities of users, combined with the interactive visualization of Spotfire
  ◦ Integrate R, S+, SAS, and MATLAB into Spotfire and custom applications
  ◦ Enable your users to utilize powerful analytic capabilities without needing a stats background
  ◦ Enable statisticians to ensure compliance and proper usage, while making their work more widely and easily available
  ◦ Utilize the analytic power of Teradata Aster for in-database predictive analytics for applications such as determining the effectiveness of content, website “golden path” analysis, and viewer engagement

• Leverage your existing analytic investments and skills to improve decision-making across your organization

• Integrate with the new, enterprise-class, R-compatible statistical engine: TIBCO Enterprise Runtime for R (TERR), which can be run locally under Spotfire Professional for offline use, or remotely through TSSS

• Tightly integrate with the Spotfire platform, as well as with open C# and Java APIs for integration of advanced analytics into custom applications

• Ensure enterprise reliability, with features such as clustering and load-balancing

• Enable users to get started rapidly using the provided OOTB (out-of-the-box) predictive analytics, and to quickly learn from templates and examples
TIBCO Spotfire full analytic application authoring

Building Spotfire applications that leverage predictive analytics is a quick and easy process. It starts with the data scientist prototyping an analytic in their environment of choice (R, SAS, etc.), and then deploying the analytic to Statistics Services. As part of this process, the data scientist specifies the types of inputs and outputs the analytic expects. This makes it available to the Spotfire application developer who, without any coding or requiring any deep understanding of the details of the analytic, uses the information provided by the data scientist to integrate the analytic into a Spotfire application. This application can then be quickly shared to a wide community of users across your organization.
In-Database Predictive Analytics via Teradata Aster

The most recent addition (as of Spotfire 5.5) to the predictive analytic ecosystem is Teradata Aster. Spotfire users can now utilize Teradata Aster to do in-database predictive analytics on Big Data from Spotfire applications or TERR scripts. They can use the analytic power of Aster for applications such as determining the effectiveness of web content, website “golden path” analysis, and viewer engagement, all without unnecessarily moving the data from the database.

This connection is implemented as a TERR package called AsterDB, which generates the SQL/Map Reduce scripts needed to access the powerful functionality within Aster. This makes it easy to leverage these in-database advanced analytics from both TERR scripts and Spotfire applications, using example templates provided with TSSS.
TIBCO Enterprise Runtime for R (TERR)

TERR is an enterprise-grade analytic engine that TIBCO built from the ground up to be fully compatible with the R language, leveraging our long-time expertise in the closely related S+ analytic engine. This allows customers to continue to develop in open source R, but to then integrate and deploy their R code on a commercially-supported and robust platform – without the need to rewrite their code.

TERR enables organizations to:

- Apply consistent models across multiple applications and uses, from prototyping to production
  - Eliminating uncertainty when analytic models implemented on disparate platforms disagree
- Easily compare multiple analytic approaches to find the hidden insights and to make the best decisions
  - And then broadly leverage these insights across the organization
- Eliminate time/resources spent re-implementing R code for production, or time spent prototyping on an unwieldy platform
  - Reducing the need for multiple analytic platforms
- Rapidly cycle from prototyping to production to deliver faster time to insight/market
  - And continually refine models and provide consistent application across the organization so that everyone is using the right, best analytic

The main technical advantages of TIBCO Enterprise Runtime for R are:

- Higher performance and far more robust memory management – so that performance is linear as larger data is analyzed
- Fully TIBCO IP, so that TERR is licensable for embedding and redistribution (unlike Open Source R, which is GPL, a particularly viral form of open source licensing)
- An engine architected as a platform for ongoing investment to ensure analytic needs can be met both now and in the future
- Broad coverage of core R functionality and CRAN packages

All these features were developed with the goal of delivering analytic power AND agility, so that customers can develop in open source R, and deploy/scale/integrate using Enterprise Runtime for R, without having to recode their analytics. People often build prototypes in R, but then typically re-implement in another language for production purposes because R was not built for enterprise usage. TERR brings enterprise-class scalability and stability to the agile R language, and enables statistics to broadly share their analyses through TIBCO Spotfire Statistics Services or by directly embedding the TERR engine.

TERR enables customers to rapidly iterate from prototyping to production without wasting time and effort recoding and retesting their analyses, allowing them to more rapidly respond to opportunities and threats, and easily integrate standardized predictive analytics consistently across organization.
TERR Developer’s Edition and TERR Community

A free Developer Edition of TERR is available through the TERR Community Site. This enables customers to test their R code prior to deployment and integration on a full-featured version of the TERR engine, freely available for non-production use. The TERR Community site also provides a forum for feedback, support, and collaboration of R/TERR users, and detailed information on topics such as TERR’s coverage of R functionality and CRAN packages.

The TERR Developer’s Edition is currently a console-only version because we expect R users to continue to develop in their R environment of choice, and then to test their code in TERR Developer’s Edition prior to deployment and integration. The TERR Community does include advice on using TERR with some popular R interfaces, such as ESS-Emacs and Notepad++.

Integration Options for TERR

TERR provides three levels of integration options:

TERR in Spotfire
- For: Ad hoc tools and interactive applications powered by advanced
- Full benefits of Spotfire Analytics platform:
  - Interactive visualization & data discovery
  - Easily build and share applications, leverage broad data access, etc.

TERR in Statistics Services
- For: Distributed analytics
  - Managed pools of engines
  - Load balancing, queuing, failover, parallelization, etc.
  - High level APIs for loose custom integration, data i/o (C#, Java)
  - Central management of analytics, R packages

Embeddable TERR Engine
- Custom (tight) integration, batch, existing grids, etc.
  - Faster than R, more robust, better memory management, fully supported
  - Low level APIs for tight integration

TERR has wide integration across the TIBCO platform, enabling customers to deploy consistent analytics across their organization:

- TERR is embedded in Spotfire Professional, where it powers the predictive modeling tools
- TERR can also be embedded in Spotfire applications, to make the power of advanced analytics available to all Spotfire users to enhance their decision-making. To enable this, it can be called locally under Spotfire Professional, or remotely through Statistics Services
- TERR can be embedded in TIBCO Business Events for complex event processing, to provide a scalable, high-throughput, low-latency analytic service for real-time applications such as fraud detection and customer scoring
- TERR can post messages to TIBCO’s enterprise social networking product, tibbr, enabling TERR/R users to collaborate with their peers, share their results to contribute to ongoing discussions, and easily send notifications of long-running analytic results
TIBCO and the R Community

As the commercial provider of S+ since the acquisition of Insightful in 2008, we are uniquely suited to contribute to the R community, building on the long history of collaboration within the joint R/S+ community. Our contributions include:

- The freely available TERR Developer Edition for non-production use by all R users
- The TERR Community, a support and collaboration forum
- As we develop new functionality in TERR, or port existing functionality from S+, the frequent release of these capabilities as CRAN packages or directly to R authors, so that R users can continue to develop in OS R, and deploy in TERR
- This includes S+SeqTrial, S+FlexBayes, and time series packages
- Benefactor of the R Foundation
- Co-sponsor of useR! conferences
- Ongoing contributor to R Core, and as we develop TERR, sharing of our observations and identified bugs with the R Core team, so that open source R development can benefit from our investment

TIBCO is committed to helping our customers leverage their investments in R and S+, and other analytic environments, so they can solve high value problems and make better decisions across their organization.

- Deployment of R/S+, as well as SAS and MATLAB scripts through TIBCO Spotfire
- Collaboration with partners building customer solutions on R and S+
- Ongoing support of S+, R’s commercial sibling, based on the same S language originally developed at Bell Labs

TIBCO strongly supports open source software (OSS) development, which sparks creativity, innovation and productivity, benefiting users, developers, and vendors. Beyond our work with the R Community, TIBCO has made major contributions to the OSS community, including releasing Ajax General Interface source code to the Dojo Foundation, and the core of the TIBCO PageBus to the OpenAjax Alliance.
**Predictive Modeling Tools**

Predictive modeling tools in Spotfire provide deep predictive insights into your data as part of ad hoc analysis. These tools support a full workflow for “real” predictive modeling and enable you to create, evaluate, and iterate predictive models while leveraging the full interactivity and powerful visualizations of the Spotfire platform. You also can test models on existing data, apply predictions to new data, and embed predictive models in applications – all without requiring any R/S+ programming.

These predictive modeling tools – Linear and Logistic Regression, Classification, and Regression Trees – are available in Spotfire Professional. They are located directly in the Spotfire menu and execute locally, using the TERR engine behind the scenes. Statistics Services is not required to use the tools for ad hoc analysis, but it does enable models developed with these tools to be deployed to applications running in the Spotfire Web Player.
Custom Advanced Analytic Tools

Spotfire is an extensible platform, so customers can also create their own, custom advanced analytic tools in the Spotfire platform, through C# coding in the Spotfire interface and through scripting their analytics. These analytics can be written in R and run locally using TERR, or written in any of the analytic engines that comprise the Spotfire Predictive Analytics ecosystem and executed remotely using Statistics Services.

This enables Spotfire users to provide custom tools highly targeted to specific users and workflows to improve their productivity, and leverage their existing investments in R scripts and other advanced analytics.
For more information
TIBCO Spotfire Website: http://spotfire.tibco.com/
TIBCO Spotfire Demo Library: http://spotfire.tibco.com/home/demos
Trends and Outliers - TIBCO Spotfire’s Business Intelligence Blog: http://spotfire.tibco.com/blog/
TERR Community Site: https://www.tibcommunity.com/community/products/analytics/terr