

IBM SPSS Modeler 18.5 评分适配器安装



内容

第 1 章 IBM SPSS Modeler 评分适配器安装.....	1
IBM SPSS Modeler 评分适配器安装.....	1
关于评分.....	1
迁移至新版本.....	1
安装用于 Netezza 的 IBM SPSS Modeler Server 评分适配器.....	1
安装用于 Teradata 的 IBM SPSS Modeler Server 评分适配器.....	2
安装 IBM SPSS Modeler Server Scoring Adapter for DB2 LUW.....	4

第 1 章 IBM SPSS Modeler 评分适配器安装

IBM SPSS Modeler 评分适配器安装

对于某些数据库，大多数 SPSS Modeler 模型块可启用 SQL 回送。通过此方法，可在数据库内执行模型评分，而不必在评分前抽取数据。此回送可使用 SPSS Modeler 中的本地 SQL，也可使用为不同数据库量身定制的附加 SQL 评分适配器（如可用）。

评分适配器支持对流中大多数模型块进行评分，以下情况除外：

- 不支持关联规则（包含列表数据）、时间序列、序列、PCA、STP 和 TCM。
- 关联模型不支持事务格式。
- Db2 LUW 中支持文本分析 (TA) - 字段方式，而其他数据库中不支持。在字段方式下使用 TA 评分适配器时，返回的行长度可能包含很多评分输出列，这会导致超出一些数据库限制。例如，可能会看到错误消息：**在 Select 或 Insert 列表中返回过多项**。有时可以在数据库中更改这些限制，如果无法更改，那么建议的方法是使用记录方式来对模型进行评分。

使用评分适配器在数据库内由生成的模型对数据进行评分以避免传输数据。SPSS Modeler 支持与 IBM 和非 IBM 数据库进行集成，并支持以更高的效率更快地部署模型。

如果评分适配器安装到相关数据库中，那么 SQL 生成选项缺省情况下生成评分适配器 SQL，除非您特别选择将其覆盖。可使用评分适配器的数据库为：

- Netezza
- Teradata
- Db2 LUW

关于评分

在 IBM® SPSS Modeler 中，评分数据被定义为使用未知结果在新的数据上部署预测模型。此预测模型处理传入数据，并且给出有关事件发生可能性或概率的预测分数。例如，当发生在线支付交易时，预测模型将处理输入数据并提供一个预测分数，给出交易是真的或是欺诈交易的概率。

SPSS Modeler 内常规的处理方法是，当预测模型接收到传入数据，它将使用数据库中的历史数据计算输入并创建预测分数的输出。此分数提供了构建预测分析模型所针对的事件的概率。

使用评分适配器的预测模型处理方法与上述方法不同的地方是，评分适配器对每个记录进行求值，并且无需从数据库中导出数据就能得到分数或预测，在模型中运行得到的分数或预测并重新导入回数据库，从而使整个处理过程更快。

迁移至新版本

如果要迁移至新版本的 IBM SPSS Modeler Server Scoring Adapter，请注意以下信息。

- 仅支持来自先前版本的适配器迁移。不支持从较旧版本迁移。例如，对于使用 Modeler Adapter 18.1.1 的 IBM SPSS 协作和部署服务 8.1.1，支持从使用 Modeler Adapter 17.1 的 IBM SPSS 协作和部署服务 7 进行迁移。
- Modeler Adapter 可与 IBM SPSS 协作和部署服务 同时安装或者在安装 IBM SPSS 协作和部署服务 后进行安装。

有关迁移的完整详细信息和指示信息，请参阅《IBM SPSS 协作和部署服务 Repository Server 安装和配置指南》的迁移部分。Modeler Adapter 的迁移过程与指南中描述的过程相同。该指南在 IBM SPSS 协作和部署服务 IBM 文档 (<https://www.ibm.com/docs/SS69YH>) 中提供。

安装用于 Netezza 的 IBM SPSS Modeler Server 评分适配器

如果您具有安装的 Netezza 的评分适配器的先前版本，那么首先必须按如下所示将其卸载：

1. 在 Netezza 评分适配器的安装目录中找到名为 Uninstall IBM SPSS Modeler Server Scoring Adapter for Netezza 的可执行文件。此文件在名为 Uninstall IBM SPSS Modeler Server Scoring Adapter for Netezza 的文件夹中。
2. 从控制台或通过使用图形用户界面 (GUI) 运行可执行文件。执行卸载程序中的指示信息以卸载评分适配器。
3. 如果收到一条指示无法移除一些项的消息，请转至适配器所在的根目录（例如，cfscoring 目录），并针对列出的未移除目录运行命令 `rm -rf`。将会将它们移除。
4. 继续以下步骤以安装评分适配器的新版本。

根据您的数据库配置，即可通过控制台进行安装，也可使用图形用户界面 (GUI) 进行安装，但是这两种方法的第一步相同。

- 运行 `install.bin` 安装脚本。确保 `nz` 用户可执行 `install.bin`，然后以此用户运行。

控制台安装

1. 将显示详细的介绍信息。按 Enter 键继续。
2. 将显示许可信息。阅读许可条款，键入 Y 接受这些条款，并按 Enter 继续。
3. 将提示您输入安装位置。将显示缺省安装位置；但是，如果您的安装位于其他位置，那么请键入此位置并按 Enter 键。
注：安装必须位于 `/nz/export/` 路径下。
4. 将提示您输入数据库名称、数据库用户名和数据库密码。
注：数据库用户必须具有数据库访问许可权才能初始化数据库和注册 `udf` 模块。
5. 这将显示安装前摘要，以确认到目前为止输入的信息。按 Enter 键继续。
6. 将显示一条消息，提示安装程序已准备好开始运行。按 Enter 键继续。
7. 在安装程序运行期间，将出现一个进度条。安装完成时，请按 Enter 键以退出安装程序。

GUI 安装

1. 将显示详细的介绍信息。单击下一步以继续。
2. 将显示许可信息。阅读许可条款，选择相应选项以接受它，然后单击下一步继续。
3. 将提示您选择安装位置。将显示缺省安装位置；但是，如果您的安装位于其他位置，那么请单击选择以浏览至该位置。当显示正确的位置时，单击下一步。
4. 将提示您输入数据库名称、数据库用户名和数据库密码。
注：数据库用户必须具有数据库访问许可权才能初始化数据库和注册 `UDF` 模块。
5. 这将显示安装前摘要，以确认到目前为止输入的信息。单击安装继续。
6. 在安装程序运行期间，将出现一个进度条。安装完成后，请单击完成以退出安装程序。

在完成这些步骤之后，评分适配器就可以开始工作了。

注：Netezza UDF 具有评分适配器可处理的 64 字段限制。如果超出此字段限制，那么会显示验证错误消息 `SQL 验证错误：HY000[46] 错误：无法将超过 64 个参数传递给函数`，将在不使用评分适配器的情况下继续模型评分。

安装用于 Teradata 的 IBM SPSS Modeler Server 评分适配器

如果您具有安装的 Teradata 的评分适配器的先前版本，那么首先必须按如下所示将其卸载：

1. 在 Teradata 评分适配器的安装目录中找到名为 Uninstall IBM SPSS Modeler Server Scoring Adapter for Teradata 的可执行文件。此文件在名为 Uninstall IBM SPSS Modeler Server Scoring Adapter for Teradata 的文件夹中。
2. 从控制台或通过使用图形用户界面 (GUI) 运行可执行文件。执行卸载程序中的指示信息以卸载评分适配器。

3. 如果收到一条指示无法移除一些项的消息，请转至适配器所在的根目录（例如，`cfscoring` 目录），并针对列出的未移除目录运行命令 `rm -rf`。将会将它们移除。
4. 继续以下步骤以安装评分适配器的新版本。

根据您的数据库配置，即可通过控制台进行安装，也可使用图形用户界面 (GUI) 进行安装，但是这两种方法的第一步相同。

- 以 `root` 或 `DBA` 用户登陆，然后运行 `install.bin` 安装脚本。您必须具有安装文件夹的访问许可权才能执行此操作。安装用户还必须具有 `CREATE FUNCTION` 许可权。

控制台安装

1. 将显示详细的介绍信息。按 `Enter` 键继续。
2. 将显示许可信息。阅读许可条款，键入 `Y` 接受这些条款，并按 `Enter` 继续。
3. 将提示您输入安装位置。将显示缺省安装位置；但是，如果您的安装位于其他位置，那么请键入此位置并按 `Enter` 键。
4. 输入数据库 TDPID。按 `Enter` 键继续。
5. 输入用户名。按 `Enter` 键继续。
6. 输入密码。按 `Enter` 键继续。
7. 这将显示安装前摘要，以确认到目前为止输入的信息。按 `Enter` 键继续。
8. 将显示一条消息，提示安装程序已准备好开始运行。按 `Enter` 键继续。
9. 在安装程序运行期间，将出现一个进度条。安装完成时，请按 `Enter` 键以退出安装程序。
10. 如果数据库中存在组件表，那么将显示一条确认消息。请输入 `Y` 以继续在数据库中创建表和函数，或输入 `N` 以跳过此步骤。注：如果跳过此步骤，那么稍后必须使用 `<installation path>\setup` 文件夹中存储的 `initdb.sh` 来手动创建表和函数。

GUI 安装

1. 将显示详细的介绍信息。单击**下一步**以继续。
2. 将显示许可信息。阅读许可条款，选择相应选项以接受它，然后单击**下一步**继续。
3. 将提示您选择安装位置。将显示缺省安装位置；但是，如果您的安装位于其他位置，那么请单击**选择**以浏览至该位置。当显示正确的位置时，单击**下一步**。
4. 输入数据库 TDPID、名称和密码，然后单击**下一步**继续。
5. 这将显示安装前摘要，以确认到目前为止输入的信息。单击**安装继续**。
6. 在安装程序运行期间，将出现一个进度条。安装完成后，请单击**完成**以退出安装程序。
7. 如果数据库中存在组件表，那么将显示一条确认消息。请单击**是**以继续在数据库中创建表和函数，或单击**否**以跳过此步骤。注：如果跳过此步骤，那么稍后必须使用 `<installation path>\setup` 文件夹中存储的 `initdb.sh` 来手动创建表和函数。

在完成这些步骤之后，评分适配器就可以开始工作了。

注：将 UDF 和 COMPONENTS 表安装到安装评分适配器的用户的缺省数据库。

共享评分适配器

要与其他 Teradata 用户共享评分适配器：

1. 授予用户以下特权：
 - 安装评分适配器的数据库上的 `SELECT` 和 `EXECUTE FUNCTION`。
 - 安装评分适配器的数据库上的 `COMPONENTS` 表的 `INSERT`。
2. 在使用安装的评分适配器针对 Teradata 生成数据库连接时，打开“数据库预置”对话框，启用**使用服务器评分适配器模式**并从**服务器评分适配器模式**下拉列表中选择模式。

注：“数据库预置”对话框根据数据库的不同而不同，并且脚本编制不支持；因此，只能在 SPSS Modeler 客户机中完成此步骤。

防止日期或时间条目出现 SQL 错误

如果“日期或时间”格式的 ODBC 驱动程序设置为“整数”，并且输入表包含编码为“日期”、“时间”或“时间戳记”的字段；那么 Teradata 会显示一条 SQL 错误消息，并且无法处理这些字段。

要防止发生此错误，请按以下步骤进行操作：

1. 打开 ODBC 数据源管理器。
2. 打开使用 Teradata 驱动程序的 DSN。
3. 单击选项 >> 以打开 **Teradata ODBC 驱动程序选项** 对话框。
4. 在此对话框的右上角，将日期时间格式设置为 AAA。
5. 保存更改。
6. 在 IBM SPSS Modeler Server 中，除去连接，然后重新连接到使用 Teradata 驱动程序的 DSN。

注：在 Unix/Linux 上，此选项称为：**DateTimeFormat**。

Teradata 驱动程序和空的 datetime_now 值

当使用 DataDirect Teradata 驱动程序时，SQL_COLUMN_TYPE 可能返回空值。替代方法是使用本机 Teradata 驱动程序。

安装 IBM SPSS Modeler Server Scoring Adapter for DB2 LUW

注：Db2 LUW 评分适配器仅在运行 LINUX 或 AIX 的 Db2 上可用。

注：IBM SPSS Modeler Server Scoring Adapter 可能与 SAS 内嵌进程 DB2 LUW ANALYZE_TABLE 相冲突，因为它们共享同一个 DB2 内建支持。

如果您具有安装的 Db2 LUW 的评分适配器的先前版本，那么首先必须按如下所示将其卸载：

1. 在 Db2 评分适配器的安装目录中找到名为 Uninstall IBM SPSS Modeler Server Scoring Adapter for DB2 的可执行文件。此文件在名为 Uninstall IBM SPSS Modeler Server Scoring Adapter for DB2 的文件夹中。
2. 从控制台或通过使用图形用户界面 (GUI) 运行可执行文件。执行卸载程序中的指示信息以卸载评分适配器。
3. 如果收到一条指示无法移除一些项的消息，请转至适配器所在的根目录（例如，cfscoring 目录），并针对列出的未移除目录运行命令 `rm -rf`。将会将它们移除。
4. 继续以下步骤以安装评分适配器的新版本。

在安装之前，您需要使用 IBM SPSS Modeler Server 评分适配器安装文件夹中提供的 `db2ida_epspss.sh` 脚本来关闭 Db2 LUW ANALYZE_TABLE 嵌入式进程。要执行此操作，请使用 Db2 命令：
`db2ida_epspss.sh stop`。

安装后，应该会使用 `db2start` 自动启动 DB2 LUW ANALYZE_TABLE 内嵌进程。但是，如果您是在 DB2 实例处于活动状态时安装适配器的，那么可以使用命令“`db2ida_epspss.sh start`”手动启动 DB2 LUW ANALYZE_TABLE 内嵌进程。

注：如果 Db2 LUW ANALYZE_TABLE 包含“UNION ALL”，那么它将不支持“WITH”表表达式子句。当尝试在生成 UNION 或 UNION ALL SQL 的 IBM SPSS Modeler 节点（比如 Append 节点）中使用此表达式时，可能会发生问题。

根据您的数据库配置，即可通过控制台进行安装，也可使用图形用户界面 (GUI) 进行安装，但是这两种方法的第一步相同。

- 运行 `install.bin` 安装脚本。确保 `install.bin` 可由 **db2 用户** 执行，并以该用户身份运行。

控制台安装

1. 将显示详细的介绍信息。按 Enter 键继续。
2. 将显示许可信息。阅读许可信息，输入 1 接收该许可信息并按下 Enter 键。
3. 将提示您输入安装位置。将显示缺省安装位置；但是，如果您的安装位于其他位置，那么请键入此位置并按 Enter 键。
4. 将提示您输入数据库名称、数据库用户名和数据库密码。
5. 这将显示安装前摘要，以确认到目前为止输入的信息。按 Enter 键继续。
6. 将显示一条消息，提示安装程序已准备好开始运行。按 Enter 键继续。
7. 在安装程序运行期间，将出现一个进度条。安装完成时，请按 Enter 键以退出安装程序。

GUI 安装

1. 将显示详细的介绍信息。单击**下一步**以继续。
2. 将显示许可信息。阅读许可条款，选择相应选项以接受它，然后单击**下一步**继续。
3. 将提示您选择安装位置。将显示缺省安装位置；但是，如果您的安装位于其他位置，那么请单击**选择**以浏览至该位置。当显示正确的位置时，单击**下一步**。
4. 将提示您输入数据库名称、数据库用户名和数据库密码。
5. 这将显示安装前摘要，以确认到目前为止输入的信息。单击**安装**继续。
6. 在安装程序运行期间，将出现一个进度条。安装完成后，请单击**完成**以退出安装程序。

在完成这些步骤之后，评分适配器就可以开始工作了。

注：如果通过 Database Scoring Adapters for Db2 LUW 对大的文本挖掘模型进行评分有问题，那么您可能需要修改数据库表列大小参数。如果遇到与将模型插入到组件表中失败相关的错误，请根据情况使用诸如以下的 Db2 命令增大列大小参数：

```
ALTER TABLE COMPONENTS ALTER COLUMN MODELDS2 SET DATA TYPE BLOB(48M);
```


