

IBM SPSS Modeler 18.1.1 Scoring Adapter Installation

IBM

Inhaltsverzeichnis

Installation von IBM SPSS Modeler Scoring Adapter	1
Installation von IBM SPSS Modeler Scoring Adapter	1
Scoring	1
Installieren von IBM SPSS Modeler Server Scoring Adapter for Netezza	2

Installieren von IBM SPSS Modeler Server Scoring Adapter for Teradata	3
Installieren von IBM SPSS Modeler Server-Scoring-Adapter für Db2 LUW	5

Installation von IBM SPSS Modeler Scoring Adapter

Installation von IBM SPSS Modeler Scoring Adapter

Bei einigen Datenbanken kann für die meisten SPSS Modeler-Modellnuggets SQL-Pushback aktiviert werden. Auf diese Weise kann Modellscoring innerhalb der Datenbank durchgeführt werden, ohne dass die Daten vor dem Scoring extrahiert werden müssen. Dieses Pushback kann entweder das native SQL innerhalb von SPSS Modeler oder, falls verfügbar, zusätzliche, für unterschiedliche Datenbanken zugeschnittene SQL-Scoring-Adapter verwenden.

Die Scoring-Adapter unterstützen mit folgenden Ausnahmen das Scoring der meisten Modellnuggets in einem Stream:

- Assoziationsregeln (mit Listendaten), Zeitreihen, Sequenz, PCA, STP und TCM werden nicht unterstützt.
- Assoziationsmodelle unterstützen das Transaktionsformat nicht.
- Text Analytics (TA) - Der Feldmodus wird in Db2 LUW unterstützt, jedoch nicht in den anderen Datenbanken. Wenn der TA-Scoring-Adapter im Feldmodus verwendet wird, kann die zurückgegebene Zeilenlänge zahlreiche Scoring-Ausgabespalten enthalten, was dazu führen kann, dass einige Datenbankgrenzwerte überschritten werden. Es kann z. B. die Fehlermeldung TOO MANY ITEMS RETURNED IN SELECT OR INSERT LIST angezeigt werden. Manchmal können diese Grenzwerte in der Datenbank geändert werden. Wenn dies jedoch nicht möglich ist, wird empfohlen, das Modell im Datensatzmodus zu scoren.

Durch die Verwendung von Scoring-Adaptoren, die das Scoring von Daten durch die generierten Modelle in der Datenbank ermöglichen, wird zusätzliche Datenübertragung vermieden. SPSS Modeler ermöglicht die Integration in IBM Datenbanken sowie in Datenbanken anderer Anbieter sowie die schnellere und effizientere Bereitstellung von Modellen.

Wenn Scoring-Adapter in den entsprechenden Datenbanken installiert sind, generiert die Option für die SQL-Generierung standardmäßig Scoring-Adapter-SQL, es sei denn, Sie setzen dies ausdrücklich außer Kraft. Scoring-Adapter sind für folgende Datenbanken verfügbar:

- Netezza
- Teradata
- Db2 LUW

Scoring

In IBM® SPSS Modeler wird das Scoring von Daten als Bereitstellung eines Vorhersagemodells für neue Daten mit unbekanntem Ergebnis definiert. Dieses Vorhersagemodell verarbeitet eingehende Daten und erstellt einen Vorhersagescore zur Wahrscheinlichkeit eines Ereignisses. Bei einer Onlinezahlungstransaktion verarbeitet ein Vorhersagemodell beispielsweise die eingehenden Daten und erstellt einen Vorhersagescore, der die Wahrscheinlichkeit angibt, dass eine Transaktion echt oder Betrug ist.

Wenn in SPSS Modeler ein Vorhersagemodell eingehende Daten empfängt, wertet es die Daten normalerweise mithilfe von historischen Daten aus einer Datenbank aus und erstellt als Ausgabe einen vorhergesagten Score. Dieser Score gibt die Wahrscheinlichkeit für ein Ereignis an, für das ein Vorhersageanalysemodell erstellt wird.

Der Unterschied beim Vorhersagemodellprozess mithilfe eines Scoring-Adapters liegt darin, dass der Scoring-Adapter die Auswertung jedes Datensatzes ermöglicht und ein Score bzw. eine Vorhersage in der Datenbank erstellt wird, ohne dass die Daten aus der Datenbank exportiert, über das Modell ausgeführt und wieder zurück importiert werden müssen, wodurch der gesamte Prozess beschleunigt wird.

Installieren von IBM SPSS Modeler Server Scoring Adapter for Netezza

Wenn Sie eine Vorgängerversion des Scoring-Adapters für Netezza installiert haben, müssen Sie diese zuerst wie folgt deinstallieren:

1. Suchen Sie die ausführbare Datei `Uninstall IBM SPSS Modeler Server Scoring Adapter for Netezza` im Installationsverzeichnis des Scoring-Adapters für Netezza. Sie befindet sich im Ordner `Uninstall IBM SPSS Modeler Server Scoring Adapter for Netezza`.
2. Führen Sie die ausführbare Datei über die Konsole oder über die grafische Benutzerschnittstelle (GUI - Graphical User Interface) aus. Befolgen Sie die Anweisungen im Deinstallationsprogramm, um den Scoring-Adapter zu deinstallieren.
3. Wenn Sie eine Nachricht erhalten, dass einige Elemente nicht entfernt wurden, wechseln Sie in das Stammverzeichnis, in dem der Adapter sich befand (beispielsweise das Verzeichnis `cfscoring`) und führen Sie den Befehl `rm -rf` für die aufgeführten Verzeichnisse aus, die bisher nicht entfernt worden wurden. Dadurch werden diese Verzeichnisse entfernt.
4. Führen Sie die folgenden Schritte aus, um die neue Version des Scoring-Adapters zu installieren.

Abhängig von der Konfiguration Ihrer Datenbank können Sie über die Konsole oder über die grafische Benutzerschnittstelle (GUI - Graphical User Interface) installieren. Der erste Schritt ist jedoch bei beiden Methoden identisch:

- Führen Sie das Installationsscript `install.bin` aus. Stellen Sie sicher, dass `install.bin` vom Benutzer `nz` ausgeführt werden kann und führen Sie das Script als dieser Benutzer aus.

Konsoleninstallation

1. Einführungsdetails werden angezeigt. Drücken Sie zum Fortfahren die Eingabetaste.
2. Die Lizenzinformationen werden angezeigt. Lesen Sie die Lizenz, geben Sie `Y` ein, um sie zu akzeptieren, und drücken Sie zum Fortfahren die Eingabetaste.
3. Sie werden aufgefordert, die Installationsposition einzugeben. Die Standardinstallationsposition wird angezeigt. Wenn Ihre Installation eine andere Position verwendet, geben Sie diese ein und drücken die Eingabetaste.

Anmerkung: Die Installation muss im Pfad `/nz/export/` ausgeführt werden.

4. Sie werden aufgefordert, den Datenbanknamen, den Namen des Datenbankbenutzers und das Kennwort für die Datenbank anzugeben.

Anmerkung: Der Datenbankbenutzer muss über Zugriffsberechtigungen für die Datenbank verfügen, um die Datenbank initialisieren und die UDF-Module registrieren zu können.

5. Eine Vorinstallationsübersicht wird angezeigt, damit Sie Ihre bisherigen Eingaben bestätigen können. Drücken Sie zum Fortfahren die Eingabetaste.
6. Eine Nachricht wird angezeigt, die besagt, dass die Installationsroutine nun ausgeführt werden kann. Drücken Sie zum Fortfahren die Eingabetaste.
7. Während die Installationsroutine ausgeführt wird, wird eine Fortschrittsleiste angezeigt. Drücken Sie nach Abschluss der Installation die Eingabetaste, um das Installationsprogramm zu beenden.

GUI-Installation

1. Einführungsdetails werden angezeigt. Klicken Sie auf **Next**, um fortzufahren.
2. Die Lizenzinformationen werden angezeigt. Lesen Sie die Lizenz, wählen Sie die Option zum Akzeptieren der Lizenz aus und klicken Sie auf **Next**, um fortzufahren.
3. Sie werden aufgefordert, die Installationsposition auszuwählen. Die Standardinstallationsposition wird angezeigt. Wenn Ihre Installation eine andere Position verwendet, klicken Sie auf **Choose**, um zu dieser Position zu navigieren. Wenn die korrekte Position angezeigt wird, klicken Sie auf **Next**.
4. Sie werden aufgefordert, den Datenbanknamen, den Namen des Datenbankbenutzers und das Kennwort für die Datenbank anzugeben.

Anmerkung: Der Datenbankbenutzer muss über Zugriffsberechtigungen für die Datenbank verfügen, um die Datenbank initialisieren und die UDF-Module registrieren zu können.

5. Eine Vorinstallationsübersicht wird angezeigt, damit Sie Ihre bisherigen Eingaben bestätigen können. Klicken Sie auf **Install**, um fortzufahren.
6. Während die Installationsroutine ausgeführt wird, wird eine Fortschrittsleiste angezeigt. Klicken Sie nach Abschluss der Installation auf **Done**, um das Installationsprogramm zu beenden.

Wenn Sie diese Schritte ausgeführt haben, ist der Scoring-Adapter betriebsbereit.

Anmerkung: Für die Netezza-UDF gilt ein Feldgrenzwert von 64 Argumenten, die vom Scoring-Adapter verarbeitet werden können. Wenn dieser Feldgrenzwert überschritten wird, wird eine Validierungsfehlermeldung (HY000[46]) angezeigt, dass maximal 64 Argumente an eine Funktion übergeben werden können, und das Modellscoring wird ohne Verwendung des Scoring-Adapters fortgesetzt.

Installieren von IBM SPSS Modeler Server Scoring Adapter for Teradata

Wenn Sie eine Vorgängerversion des Scoring-Adapters für Teradata installiert haben, müssen Sie diese zuerst wie folgt deinstallieren:

1. Suchen Sie die ausführbare Datei `Uninstall IBM SPSS Modeler Server Scoring Adapter for Teradata` im Installationsverzeichnis des Scoring-Adapters für Teradata. Sie befindet sich im Ordner `Uninstall IBM SPSS Modeler Server Scoring Adapter for Teradata`.
2. Führen Sie die ausführbare Datei über die Konsole oder über die grafische Benutzerschnittstelle (GUI - Graphical User Interface) aus. Befolgen Sie die Anweisungen im Deinstallationsprogramm, um den Scoring-Adapter zu deinstallieren.
3. Wenn Sie eine Nachricht erhalten, dass einige Elemente nicht entfernt wurden, wechseln Sie in das Stammverzeichnis, in dem der Adapter sich befand (beispielsweise das Verzeichnis `cfscoring`) und führen Sie den Befehl `rm -rf` für die aufgeführten Verzeichnisse aus, die bisher nicht entfernt worden wurden. Dadurch werden diese Verzeichnisse entfernt.
4. Führen Sie die folgenden Schritte aus, um die neue Version des Scoring-Adapters zu installieren.

Abhängig von der Konfiguration Ihrer Datenbank können Sie über die Konsole oder über die grafische Benutzerschnittstelle (GUI - Graphical User Interface) installieren. Der erste Schritt ist jedoch bei beiden Methoden identisch:

- Melden Sie sich entweder als *Root* oder als Benutzer *DBA* an und führen Sie das Installationsscript `install.bin` aus. Sie müssen dazu über Zugriffsberechtigungen für das Installationsverzeichnis verfügen. Der installierende Benutzer muss darüber hinaus über die Berechtigungen `CREATE FUNCTION` verfügen.

Konsoleninstallation

1. Einführungsdetails werden angezeigt. Drücken Sie zum Fortfahren die Eingabetaste.
2. Die Lizenzinformationen werden angezeigt. Lesen Sie die Lizenz, geben Sie `Y` ein, um sie zu akzeptieren, und drücken Sie zum Fortfahren die Eingabetaste.
3. Sie werden aufgefordert, die Installationsposition einzugeben. Die Standardinstallationsposition wird angezeigt. Wenn Ihre Installation eine andere Position verwendet, geben Sie diese ein und drücken die Eingabetaste.
4. Geben Sie die TDP-ID der Datenbank ein. Drücken Sie zum Fortfahren die Eingabetaste.
5. Geben Sie den Benutzernamen ein. Drücken Sie zum Fortfahren die Eingabetaste.
6. Geben Sie das Kennwort ein. Drücken Sie zum Fortfahren die Eingabetaste.
7. Eine Vorinstallationsübersicht wird angezeigt, damit Sie Ihre bisherigen Eingaben bestätigen können. Drücken Sie zum Fortfahren die Eingabetaste.
8. Eine Nachricht wird angezeigt, die besagt, dass die Installationsroutine nun ausgeführt werden kann. Drücken Sie zum Fortfahren die Eingabetaste.

9. Während die Installationsroutine ausgeführt wird, wird eine Fortschrittsleiste angezeigt. Drücken Sie nach Abschluss der Installation die Eingabetaste, um das Installationsprogramm zu beenden.
10. Wenn die Komponententabelle (*Components*) in Ihrer Datenbank vorhanden ist, wird eine Bestätigungsnachricht angezeigt. Geben Sie **Y** ein, um die Erstellung von Tabellen und Funktionen in Ihrer Datenbank fortzusetzen, oder geben Sie **N** ein, um diesen Schritt zu überspringen. *Hinweis:* Wenn Sie diesen Schritt überspringen, müssen Sie Tabellen und Funktionen später mithilfe des Scripts `initdb.sh` erstellen, das im Ordner `<Installationspfad>\setup` gespeichert ist.

GUI-Installation

1. Einführungsdetails werden angezeigt. Klicken Sie auf **Next**, um fortzufahren.
2. Die Lizenzinformationen werden angezeigt. Lesen Sie die Lizenz, wählen Sie die Option zum Akzeptieren der Lizenz aus und klicken Sie auf **Next**, um fortzufahren.
3. Sie werden aufgefordert, die Installationsposition auszuwählen. Die Standardinstallationsposition wird angezeigt. Wenn Ihre Installation eine andere Position verwendet, klicken Sie auf **Choose**, um zu dieser Position zu navigieren. Wenn die korrekte Position angezeigt wird, klicken Sie auf **Next**.
4. Geben Sie die TDP-ID, den Namen und das Kennwort der Datenbank ein und klicken Sie auf **Next**, um fortzufahren.
5. Eine Vorinstallationsübersicht wird angezeigt, damit Sie Ihre bisherigen Eingaben bestätigen können. Klicken Sie auf **Install**, um fortzufahren.
6. Während die Installationsroutine ausgeführt wird, wird eine Fortschrittsleiste angezeigt. Klicken Sie nach Abschluss der Installation auf **Done**, um das Installationsprogramm zu beenden.
7. Wenn die Komponententabelle (*Components*) in Ihrer Datenbank vorhanden ist, wird eine Bestätigungsnachricht angezeigt. Klicken Sie auf **Yes**, um die Erstellung von Tabellen und Funktionen in Ihrer Datenbank fortzusetzen, oder auf **No**, um diesen Schritt zu überspringen. *Hinweis:* Wenn Sie diesen Schritt überspringen, müssen Sie Tabellen und Funktionen später mithilfe des Scripts `initdb.sh` erstellen, das im Ordner `<Installationspfad>\setup` gespeichert ist.

Wenn Sie diese Schritte ausgeführt haben, ist der Scoring-Adapter betriebsbereit.

Anmerkung: Die UDFs und die Komponententabelle (*Components*) werden in der Standarddatenbank des Benutzers installiert, der den Scoring-Adapter installiert.

Gemeinsame Nutzung des Scoring-Adapters

So geben Sie den Scoring-Adapter für die gemeinsame Nutzung durch andere Teradata-Benutzer frei:

1. Erteilen Sie dem Benutzer die folgenden Berechtigungen:
 - `SELECT` und `EXECUTE FUNCTION` für die Datenbank, in der der Scoring-Adapter installiert ist.
 - `INSERT` für die Komponententabelle (*Components*) in der Datenbank, in der der Scoring-Adapter installiert ist.
2. Nachdem eine Datenbankverbindung zu Teradata bei installiertem Scoring-Adapter hergestellt wurde, öffnen Sie das Dialogfeld **Datenbank-Voreinstellungen**, aktivieren **Server-Scoring-Adapter-Schema verwenden** und wählen das Schema in der Dropdown-Liste **Server-Scoring-Adapter-Schema** aus.

Anmerkung: Das Dialogfeld **Datenbank-Voreinstellungen** sieht für unterschiedliche Datenbanken unterschiedlich aus und wird für Scripting nicht unterstützt. Daher kann dieser Schritt nur in SPSS Modeler Client ausgeführt werden.

Verhindern von SQL-Fehlern für Datums- oder Zeiteinträge

Wenn der ODBC-Treiber für das Datums- oder Zeitformat auf eine Ganzzahl gesetzt ist und Ihre Eingabetabelle Felder enthält, die als Datum, Uhrzeit oder Zeitmarke codiert sind, zeigt Teradata eine SQL-Fehlernachricht an und kann diese Felder nicht verarbeiten.

Führen Sie die folgenden Schritte aus, um diesen Fehler zu verhindern:

1. Öffnen Sie den ODBC-Datenquellenadministrator.
2. Öffnen Sie den DSN, der den Teradata-Treiber verwendet.
3. Klicken Sie auf **Optionen >>**, um das Dialogfeld mit Optionen für den Teradata-ODBC-Treiber zu öffnen.
4. Setzen Sie oben rechts im Dialogfeld das Datum-/Zeitformat auf AAA.
5. Speichern Sie Ihre Änderungen.
6. Entfernen Sie in IBM SPSS Modeler Server die Verbindung und stellen Sie sie erneut zu dem DSN her, der den Teradata-Treiber verwendet.

Anmerkung: Unter UNIX/Linux lautet diese Option **DateTimeFormat**.

Teradata-Treiber und datetime_now-Nullwerte

Wenn Sie den DataDirect Teradata-Treiber verwenden, gibt SQL_COLUMN_TYPE möglicherweise einen Nullwert zurück. Alternativ können Sie den nativen Teradata-Treiber verwenden.

Installieren von IBM SPSS Modeler Server-Scoring-Adapter für Db2 LUW

Anmerkung: Der Db2 LUW-Scoring-Adapter ist nur für Db2 unter LINUX oder AIX verfügbar.

Anmerkung: IBM SPSS Modeler Server Scoring Adapter kann mit dem integrierten Db2 LUW-Prozess ANALYZE_TABLE für SAS einen Konflikt verursachen, da beide dieselbe integrierte Db2-Unterstützung nutzen.

Wenn Sie eine Vorgängerversion des Scoring-Adapters für Db2 LUW installiert haben, müssen Sie diese zuerst wie folgt deinstallieren:

1. Suchen Sie die ausführbare Datei Uninstall IBM SPSS Modeler Server Scoring Adapter for Db2 im Installationsverzeichnis des Scoring-Adapters für Db2. Sie befindet sich im Ordner Uninstall IBM SPSS Modeler Server Scoring Adapter for Db2.
2. Führen Sie die ausführbare Datei über die Konsole oder über die grafische Benutzerschnittstelle (GUI - Graphical User Interface) aus. Befolgen Sie die Anweisungen im Deinstallationsprogramm, um den Scoring-Adapter zu deinstallieren.
3. Wenn Sie eine Nachricht erhalten, dass einige Elemente nicht entfernt wurden, wechseln Sie in das Stammverzeichnis, in dem der Adapter sich befand (beispielsweise das Verzeichnis cfscoring) und führen Sie den Befehl `rm -rf` für die aufgeführten Verzeichnisse aus, die bisher nicht entfernt worden wurden. Dadurch werden diese Verzeichnisse entfernt.
4. Führen Sie die folgenden Schritte aus, um die neue Version des Scoring-Adapters zu installieren.

Vor der Installation müssen Sie den integrierten Db2 LUW-Prozess ANALYZE_TABLE mithilfe des Scripts `db2ida_epsps.sh` beenden, das im Installationsordner des Scoring-Adapters von IBM SPSS Modeler Server bereitgestellt wird. Verwenden Sie hierzu den folgenden Db2-Befehl: `db2ida_epsps.sh stop`.

Nach der Installation sollte der integrierte Db2 LUW-Prozess ANALYZE_TABLE automatisch mithilfe von `db2start` gestartet werden. Wenn Sie jedoch den Adapter installieren, während eine Db2-Instanz aktiv ist, können Sie den integrierten Db2 LUW-Prozess ANALYZE_TABLE mithilfe des folgenden Befehls manuell starten: `db2ida_epsps.sh start`.

Anmerkung: ANALYZE_TABLE von Db2 LUW unterstützt die Klausel "WITH Tabellenausdruck" nicht, wenn sie UNION ALL enthält. Dies kann zu einem Fehler führen, wenn Sie versuchen, diesen Ausdruck in einem IBM SPSS Modeler-Knoten zu verwenden, der UNION- oder UNION ALL-SQL generiert, wie beispielsweise der Anhangknoten.

Abhängig von der Konfiguration Ihrer Datenbank können Sie über die Konsole oder über die grafische Benutzerschnittstelle (GUI - Graphical User Interface) installieren. Der erste Schritt ist jedoch bei beiden Methoden identisch:

- Führen Sie das Installationsscript `install.bin` aus. Stellen Sie sicher, dass `install.bin` vom Benutzer **db2** ausgeführt werden kann und führen Sie das Script als dieser Benutzer aus.

Konsoleninstallation

1. Einführungsdetails werden angezeigt. Drücken Sie zum Fortfahren die Eingabetaste.
2. Die Lizenzinformationen werden angezeigt. Lesen Sie die Lizenz, geben Sie 1 ein, um sie zu akzeptieren, und drücken Sie die Eingabetaste.
3. Sie werden aufgefordert, die Installationsposition einzugeben. Die Standardinstallationsposition wird angezeigt. Wenn Ihre Installation eine andere Position verwendet, geben Sie diese ein und drücken die Eingabetaste.
4. Sie werden aufgefordert, den Datenbanknamen, den Namen des Datenbankbenutzers und das Kennwort für die Datenbank anzugeben.
5. Eine Vorinstallationsübersicht wird angezeigt, damit Sie Ihre bisherigen Eingaben bestätigen können. Drücken Sie zum Fortfahren die Eingabetaste.
6. Eine Nachricht wird angezeigt, die besagt, dass die Installationsroutine nun ausgeführt werden kann. Drücken Sie zum Fortfahren die Eingabetaste.
7. Während die Installationsroutine ausgeführt wird, wird eine Fortschrittsleiste angezeigt. Drücken Sie nach Abschluss der Installation die Eingabetaste, um das Installationsprogramm zu beenden.

GUI-Installation

1. Einführungsdetails werden angezeigt. Klicken Sie auf **Next**, um fortzufahren.
2. Die Lizenzinformationen werden angezeigt. Lesen Sie die Lizenz, wählen Sie die Option zum Akzeptieren der Lizenz aus und klicken Sie auf **Next**, um fortzufahren.
3. Sie werden aufgefordert, die Installationsposition auszuwählen. Die Standardinstallationsposition wird angezeigt. Wenn Ihre Installation eine andere Position verwendet, klicken Sie auf **Choose**, um zu dieser Position zu navigieren. Wenn die korrekte Position angezeigt wird, klicken Sie auf **Next**.
4. Sie werden aufgefordert, den Datenbanknamen, den Namen des Datenbankbenutzers und das Kennwort für die Datenbank anzugeben.
5. Eine Vorinstallationsübersicht wird angezeigt, damit Sie Ihre bisherigen Eingaben bestätigen können. Klicken Sie auf **Install**, um fortzufahren.
6. Während die Installationsroutine ausgeführt wird, wird eine Fortschrittsleiste angezeigt. Klicken Sie nach Abschluss der Installation auf **Done**, um das Installationsprogramm zu beenden.

Wenn Sie diese Schritte ausgeführt haben, ist der Scoring-Adapter betriebsbereit.

Anmerkung: Wenn beim Scoring großer Textmining-Modelle über die Datenbankscoring-Adapter für Db2 LUW Probleme auftreten, müssen Sie möglicherweise den Parameter für die Spaltengröße der Datenbanktabelle ändern. Wenn beim Einfügen des Modells in die Komponententabelle Fehler auftreten, erhöhen Sie mit einem Db2-Befehl wie dem folgenden den Wert für den Parameter für die Spaltengröße entsprechend:

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




Gedruckt in Deutschland