

SQL SERVER & Machine Learning

Tomasz Krawczyk

Agenda

- › History
- › Machine Learning Server and Machine Learning Services
- › SQL Server and Machine Learning Services
- › SQL +R+ Python
 - Train models
 - Run models

History

- › MS SQL Server 2016
 - R in SQL Server 2016
- › Microsoft R Server ScaleR
- › **RevoScaleR** package
 - <https://mran.microsoft.com/packages>



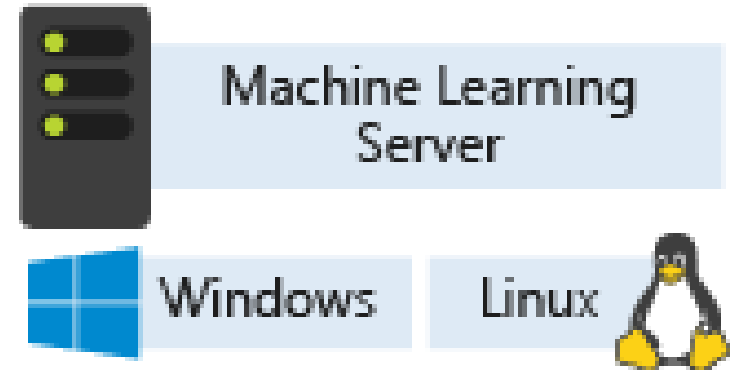
History

- › MS SQL Server 2017
 - Python in SQL Server 2017
- › Microsoft R Server -> Machine Learning Server (Services)
- › RevoScalePy
- › Microsoftml for Python



Install and configure Machine Learning Server

- › Microsoft R Server 9.1
 - R Server for Hadoop
 - R Server for Linux
 - R Server for Windows
 - R Server for Teradata
 - SQL Server
- › Machine Learning Server 9.2.1
 - Machine Learning Server for Hadoop
 - Machine Learning for Linux
 - Machine Learning for Windows
 - SQL Server



[Learn about HDInsight and cluster versions. →](#)

Cluster configuration

* Cluster type ⓘ	* Operating system	* Version
<input type="text" value="R Server"/>	<input type="text" value="Linux"/>	<input type="text" value="R Server 9.1 (HDI 3.6)"/>

R Server : Terabyte-scale, enterprise grade R analytics with transparent parallelization on top of Spark and Hadoop.

Configuration Options:

- R Server 9.1 on Spark 2.1 with Java 8
- R Server 9.0 on Spark 2.0 with Java 8
- R Server 8.0 on Spark 1.6 with Java 7

SQL Server Editions and Machine Learning Services

- › Machine Learning Services
 - Express with Advanced Services (*)
 - Developer
 - **Standard**
 - **Enterprise**

(*)

- Basic R integration
- Basic Python integration

Install Machine Learning Services on SQL Server 2017

Feature Selection

Select the Evaluation features to install.

Product Key
License Terms
Global Rules
Product Updates
Install Setup Files
Install Rules
Feature Selection
Feature Rules
Instance Configuration
Server Configuration
Database Engine Configuration
Consent to install Python
Feature Configuration Rules
Ready to Install
Installation Progress
Complete

Features:

Instance Features

- ☒ Database Engine Services
- ☐ SQL Server Replication
- ☒ Machine Learning Services (In-Database)
 - ☐ R
 - ☒ Python
- ☐ Full-Text and Semantic Extractions for Search
- ☐ Data Quality Services
- ☐ PolyBase Query Service for External Data
- ☐ Analysis Services
- ☐ Reporting Services - Native

Shared Features

- ☐ Machine Learning Server (Standalone)
 - ☐ R
 - ☐ Python
- ☐ Data Quality Client
- ☐ Client Tools Connectivity
- ☐ Integration Services
 - ☐ Scale Out Master
 - ☐ Scale Out Worker
- ☐ Client Tools Backwards Compatibility
- ☐ Client Tools SDK
- ☐ Documentation Components
- ☐ Distributed Replay Controller
- ☐ Distributed Replay Client
- ☐ SQL Client Connectivity SDK
- ☐ Master Data Services

Redistributable Features

Select All Unselect All

Instance root directory: C:\Program Files\Microsoft SQL Server\ ...

Shared feature directory: C:\Program Files\Microsoft SQL Server\ ...

Shared feature directory (x86): C:\Program Files (x86)\Microsoft SQL Server\ ...

Feature description:

The configuration and operation of each instance feature of a SQL Server instance is isolated from other SQL Server instances. SQL Server instances can operate side-by-side on the same computer.

Prerequisites for selected features:

Already installed:

- Windows PowerShell 3.0 or higher
- Microsoft .NET Framework 4.6

To be installed from media:

- Microsoft Visual Studio 2010 Redistributables
- Microsoft Visual C++ 2015 Redistributable
- Microsoft MPI v7

Disk Space Requirements

Drive C: 1904 MB required, 66869 MB available

< Back Next > Cancel

SQL Server 2017 RC2 Setup

Configuration

Service accounts and collation configuration.

Service Accounts Collation

Microsoft recommends that you use a separate account for each SQL Server service.

Service	Account Name	Password	Startup Type
SQL Server Agent	NT Service\SQLAgent\$SQL2017...		Manual
SQL Server Database Engine	NT Service\MSSQL\$SQL2017RC2		Automatic
SQL Server Launchpad	NT Service\MSSQLLaunchpad\$S...		Automatic
SQL Server Browser	NT AUTHORITY\LOCALSERVICE		Automatic

☐ Grant Perform Volume Maintenance Task privilege to SQL Server Database Engine Service

This privilege enables instant file initialization by avoiding zeroing of data pages. This may lead to information disclosure by allowing deleted content to be accessed.

[Click here for details](#)

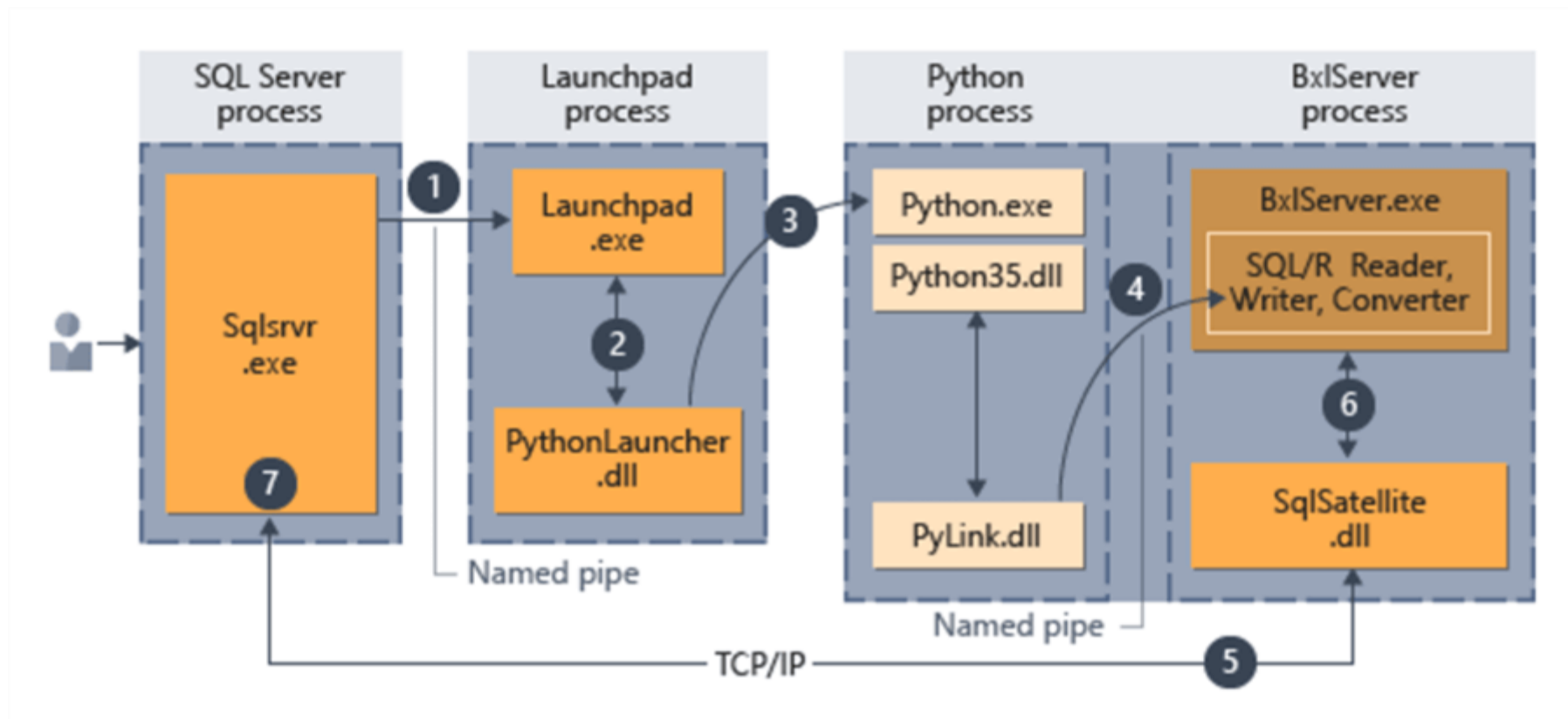
< Back Next > Cancel

Python and SQL Server Integration Architecture

› Process

- **Launchpad** - a service provided by SQL Server 2017 for supporting execution of external scripts
- **Binary Exchange Language (Bxl) Server** - manages communication between SQL Server and the Python runtime
- **SQL Satellite** - BxlServer uses SQL Satellite for communicating with SQL Server

Python and SQL Server Integration Architecture



Enable Python or (R) script execution

- › Enable External Script Execution

```
EXEC sp_configure 'external scripts enabled', 1  
RECONFIGURE WITH OVERRIDE
```

- › Restart the SQL Server service for the SQL Server instance.
- › Verify that the external script execution feature is running

```
EXEC sp_configure 'external scripts enabled'
```

How to run Python or R Script?

```
sp_execute_external_script
@language = N'language' ,
@script = N'script',

@input_data_1 = ] 'input_data_1'
[ , @input_data_1_name = ] N'input_data_1_name' ]
[ , @output_data_1_name = 'output_data_1_name' ]
[ , @parallel = 0 | 1 ]
[ , @params = ] N'@parameter_name data_type [ OUT | OUTPUT ] [ ,...n ]'
[ , @parameter1 = ] 'value1' [ OUT | OUTPUT ] [ ,...n ]
[ WITH <execute_option> ]

[;]

<execute_option>::=
{
    { RESULT SETS UNDEFINED }
    | { RESULT SETS NONE }
    | { RESULT SETS ( <result_sets_definition> ) }
}

<result_sets_definition> ::=
{
    (
        { column_name
          data_type
          [ COLLATE collation_name ]
          [ NULL | NOT NULL ] }
        [,...n ]
    )
    | AS OBJECT
      [ db_name . [ schema_name ] . | schema_name . ]
      {table_name | view_name | table_valued_function_name }
    | AS TYPE [ schema_name.]table_type_name
}
```

```
execute sp_execute_external_script
@language = N'Python',
@script = N'
import sys
print("*****")
print(sys.path)
print(sys.version)
print("Hello World")
print("*****")'
```

DEMO(s)

Demo 1,2,3
Python Env
Jupyter

Revoscale and RevoscalePy

- › The **revoscalepy** module is a collection of portable, scalable and distributable Python functions used for importing, transforming, and analyzing data at scale

<https://docs.microsoft.com/en-us/machine-learning-server/python-reference/revoscalepy/revoscalepy-package>

Microsoftml package

- › The **microsoftml** module is a collection of Python functions used in machine learning solutions. It includes functions for training and transformations, scoring, text and image analysis, and feature extraction for deriving values from existing data.

<https://docs.microsoft.com/en-us/machine-learning-server/python-reference/microsoftml/microsoftml-package>

Native Scoring

› **PREDICT** (Transact-SQL)

```
PREDICT
(
    MODEL = @model | model_literal,
    DATA = object AS <table_alias>
)
WITH ( <result_set_definition> )

<result_set_definition> ::=
{
    { column_name
      data_type
      [ COLLATE collation_name ]
      [ NULL | NOT NULL ]
    }
    [...n ]
}

MODEL = @model | model_literal
```

Native Scoring Supported Models

- › Not all models are supported - only the following models are supported:
 - rxLinMod
 - rxLogit
 - rxBTrees
 - rxDtree
 - rxdForest

π

DEMO(s) Q&A

› Resources:

<https://docs.microsoft.com/en-us/machine-learning-server/>

› Contact:

tkrawczyk@future-processing.com