Machine Learning

RESTful R Extensibility Instructions

Contents

Initial Steps 1

Repeatable Steps (for each data mart) 1

Repeatable Steps (for each destination entity) 2

Sample Seed Scripts 2

This document instructs the user how to inject their designated R scripts, served from an external RESTful server, into the Catalyst loaders. It begins by installing the required SSIS package and defining new system level attributes. It continues by injecting the new SSIS package into the designated loader step. It concludes by defining the necessary variables for each destination entity.

Initial Steps

1. Establish local folder on ETL server.
   1. Configure permissions to allow EDW loader account to read, write, and execute
2. Install ExternalScriptExtensibility ISPAC on ETL server
   1. Locate inside folder \SSISDB\CatalsytExtensibility\
   2. Verify permission to allow EDW loader account to read and execute
3. Seed new attribute names into EDWAdmin.CatalystAdmin.AttributeBASE
   * + fabric.machinelearning.server
     + fabric.machinelearning.sourceentity
     + fabric.machinelearning.bindingscript
4. Seed new system attribute value into EDWAdmin.CatalystAdminObjectAttributeBASE
   * + fabric.machinelearning.server: “http://datsci01.hqcatalyst.local:8080”

Repeatable Steps (for each data mart)

1. Seed EDWAdmin.CatalystAdmin.ETLEngineConfigurationBASE
   * + ExtensionPointNM: “OnPostStageToProdLoad”
     + DatamartID: {data mart ID}
     + SSISPackagePathOrSPToExecuteTXT: “\SSISDB\CatalystExtensibility\ExternalScriptExtensibility\ExternalScriptExecutionREST.dtsx”
     + IsSSISPackageFLG: 1
     + ExtensionOwnerNM: “Health Catalyst”
     + RunInAsyncModeFLG: 0
     + Use32BitRuntimeFLG: 0
     + ExecutionOrderNBR: 1
     + ActiveFLG: 1
     + RequiredParametersTXT: “BatchID, TableID”
     + FailsBatchFLG: 1

Repeatable Steps (for each destination entity)

1. Seed new destination entity attribute values into EDWAdmin.CatalystAdmin.ObjectAttributeBASE
   * + fabric.machinelearning.sourceentity: “SAM.Sample.SampleEntity”
     + fabric.machinelearning.bindingscript: “jobs/healthcareaitest.R”
2. Configure dependency based on need

Sample Seed Scripts

**EDWAdmin.CatalystAdmin.AttributeBASE**

IF NOT EXISTS (SELECT \* FROM CatalystAdmin.AttributeBASE WHERE AttributeNM = 'fabric.machinelearning.server')

INSERT INTO CatalystAdmin.AttributeBASE (AttributeNM, AttributeDSC) VALUES ('fabric.machinelearning.server','Path to remote machine learning server')

IF NOT EXISTS (SELECT \* FROM CatalystAdmin.AttributeBASE WHERE AttributeNM = 'fabric.machinelearning.sourceentity')

INSERT INTO CatalystAdmin.AttributeBASE (AttributeNM, AttributeDSC) VALUES ('fabric.machinelearning.sourceentity','Source entity from which to calculate predictions')

IF NOT EXISTS (SELECT \* FROM CatalystAdmin.AttributeBASE WHERE AttributeNM = 'fabric.machinelearning.bindingscript')

INSERT INTO CatalystAdmin.AttributeBASE (AttributeNM, AttributeDSC) VALUES ('fabric.machinelearning.bindingscript','HCRTools script that contains HC functions')

**EDWAdmin.CatalystAdmin.ETLEngineConfigurationBASE**

INSERT INTO CatalystAdmin.ETLEngineConfigurationBASE (ExtensionPointNM, DatamartID, SSISPackagePathOrSPToExecuteTXT, IsSSISPackageFLG, ExtensionOwnerNM, RunInAsynchModeFLG, Use32BitRuntimeFLG, ExecutionOrderNBR, ActiveFLG, RequiredParametersTXT, FailsBatchFLG)

VALUES ('OnPostStageToProdLoad', [DataMartID], '\SSISDB\CatalystExtensibility\ExternalScriptExtensibility\ExternalScriptExecutionREST.dtsx', 1, 'Health Catalyst', 0, 0, 1, 1, 'BatchID, TableID', 1)

**EDWAdmin.CatalystAdmin.ObjectAttributeBASE**

INSERT INTO CatalystAdmin.ObjectAttributeBASE (ObjectID,ObjectTypeCD,AttributeNM,AttributeTypeCD,AttributeValueTXT)

VALUES (0,'System','fabric.machinelearning.server','string','[URI]')

INSERT INTO CatalystAdmin.ObjectAttributeBASE (ObjectID,ObjectTypeCD,AttributeNM,AttributeTypeCD,AttributeValueTXT)

VALUES ([TableID],'Table','fabric.machinelearning.sourceentity','string', [DatabaseNM.SchemaNM.ViewNM]')

INSERT INTO CatalystAdmin.ObjectAttributeBASE (ObjectID,ObjectTypeCD,AttributeNM,AttributeTypeCD,AttributeValueTXT)

VALUES ([TableID],'Table','fabric.machinelearning.bindingscript','string','[ScriptName]')