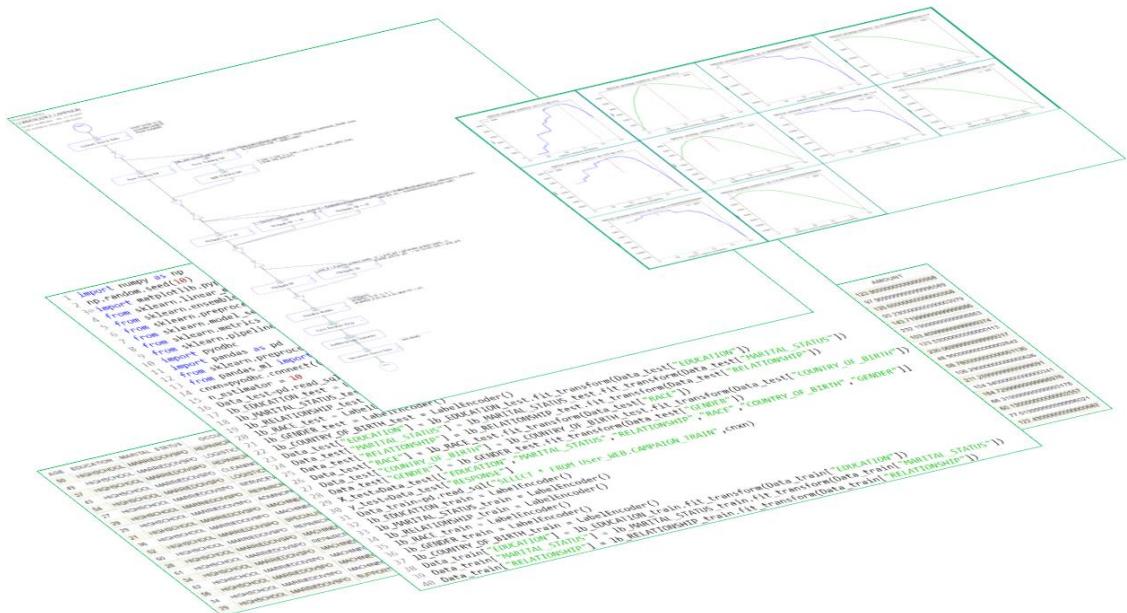


Machine Learning (ML) Toolkit

ML Toolkit Fundamentals



Document details

Title: Machine Learning (ML) Toolkit

Customer: None

Project: None

Module(s): None

Training name: ML Toolkit Fundamentals

Description: Information required for installation and use of ML Toolkit. The functionality of ML Toolkit is a prototype to be adjusted to the needs of the user. The user installs and applies ML Toolkit at their own discretion and risk.

Document ID: ML_Toolkit_Fundamentals

Version: 03

Created by: Sergey Lukyanchikov, Eduard Lebedyuk, Shiao-Bin Soong

Copyright © 2019 InterSystems Corporation. All rights reserved.

This document is confidential and proprietary. Printing renders document uncontrolled.

Document controls

Document Modifications			
Version	Date	Description of Change	Modified By
01	2019-02-19	Initial version with Python Tools	Sergey Lukyanchikov Eduard Lebedyuk
02	2019-04-11	Adding R Tools	Sergey Lukyanchikov Eduard Lebedyuk Shiao-Bin Soong
03	2019-12-18	Adding Jupyter Tools	Sergey Lukyanchikov Eduard Lebedyuk Shiao-Bin Soong

Document Authorization			
InterSystems	Name	Sergey Lukyanchikov, Eduard Lebedyuk, Shiao-Bin Soong	
	Role	Sales Engineer, Sales Engineer, Developer	
	Signature	< Insert electronic signature >	
	Date	2019-12-18	
None	Name	< Insert customer contact >	
	Role	< e.g. Project Manager >	
	Signature	< Insert electronic signature >	
	Date	< Select date >	

Contents

1. General.....	7
1.1. External Users: Open Exchange Pages	7
1.2. External Users: ML Toolkit User Group	7
1.3. Internal Users: MS Teams Group.....	8
1.4. Internal Users: OneDrive Folder.....	10
2. Python Tools.....	11
2.1. Python Gateway	11
2.2. Installation	11
2.3. Docker.....	16
2.4. Use.....	17
2.4.1. General Use.....	17
2.4.2. Shell	19
2.4.3. Context Persistence	19
2.4.4. IRIS Interoperability Adapter	19
2.4.5. Test Business Process.....	20
2.4.6. Unit Tests	21
2.4.7. ZPY Command	21
2.4.8. Limitations	21
2.5. Development	21
2.5.1. Commits	21
2.5.2. Building.....	21
2.5.3. Troubleshooting	22
2.6. Gateway Functionality.....	24
2.6.1. Execute function	24
2.6.2. Proxyless Gateway	26
2.7. Data Transfer.....	28
2.7.1. Python -> InterSystems IRIS	28
2.7.2. InterSystems IRIS -> Python	28
3. R Tools.....	32
3.1. R Gateway.....	32
3.2. Installation	32
3.3. Use.....	35
3.3.1. General Use.....	35
3.3.2. IRIS Interoperability Adapter	36
3.3.3. Notes	37
3.3.4. Sample Business Process and Production	37

3.3.5.	Unit Tests	37
3.3.6.	Limitations	37
3.3.7.	Rserve Configuration	37
3.3.8.	Development.....	40
3.3.9.	Troubleshooting	40
3.3.10.	Gateway Functionality.....	40
4.	IRIS Interoperability Tools.....	42
4.1.	Installation	42
4.2.	Use.....	44
4.2.1.	General Use.....	44
4.2.2.	Python Gateway.....	45
4.2.3.	R Gateway	51
5.	Design-Phase Tools.....	56
5.1.	Jupyter Notebook.....	56
5.1.1.	General.....	56
5.1.2.	Screenshots.....	56
5.1.3.	Installation	56
5.1.4.	Notes	57
6.	Convergent Analytics Showcases	58
6.1.	Featured Showcases: Robotized ML.....	58
6.1.1.	Background.....	58
6.1.2.	Implementation	58
6.1.3.	Walkthrough.....	59
6.2.	Featured Showcases: Robotized Sentiment Analysis (English).....	61
6.2.1.	Background.....	61
6.2.2.	Implementation	61
6.2.3.	Walkthrough.....	62
6.3.	Featured Showcases: Robotized Sentiment Analysis (Russian)	64
6.3.1.	Background.....	64
6.3.2.	Implementation	64
6.3.3.	Walkthrough.....	65
6.4.	001 Sentiment Analysis.....	67
6.4.1.	Background.....	67
6.4.2.	Implementation	68
6.4.3.	Walkthrough.....	68
6.5.	002 Engine Condition Classification	70
6.5.1.	Background.....	70
6.5.2.	Implementation	71

6.5.3.	Walkthrough.....	71
6.6.	003 Reimbursement Request Check.....	73
6.6.1.	Background.....	73
6.6.2.	Implementation	74
6.6.3.	Walkthrough.....	74
6.7.	004 Retail Cannibalization Analysis.....	76
6.7.1.	Background.....	76
6.7.2.	Implementation	77
6.7.3.	Walkthrough.....	77
6.8.	005 Marketing Campaign Optimization.....	78
6.8.1.	Background.....	78
6.8.2.	Implementation	79
6.8.3.	Walkthrough.....	80
6.9.	006 Rail Time Series Discovery	81
6.9.1.	Background.....	81
6.9.2.	Implementation	82
6.9.3.	Walkthrough.....	83
6.10.	007 Housing Debts Prediction.....	84
6.10.1.	Background.....	84
6.10.2.	Implementation	84
6.10.3.	Walkthrough.....	85
6.11.	008 Diseases Network Analysis	86
6.11.1.	Background.....	86
6.11.2.	Implementation	87
6.11.3.	Walkthrough.....	88
6.12.	009 Bell State Automation.....	89
6.12.1.	Background.....	89
6.12.2.	Implementation	89
6.12.3.	Walkthrough.....	89

1. General

1.1. External Users: Open Exchange Pages

[Open Exchange](#) is a public catalog maintained by InterSystems for a free download of functional extensions for its standard products. Those extensions are a developer community initiative with respective licenses and disclaimers specified. The following toolsets of ML Toolkit are available via Open Exchange as of the publish date of this document:

- [Python Gateway](#)



- [R Gateway](#)

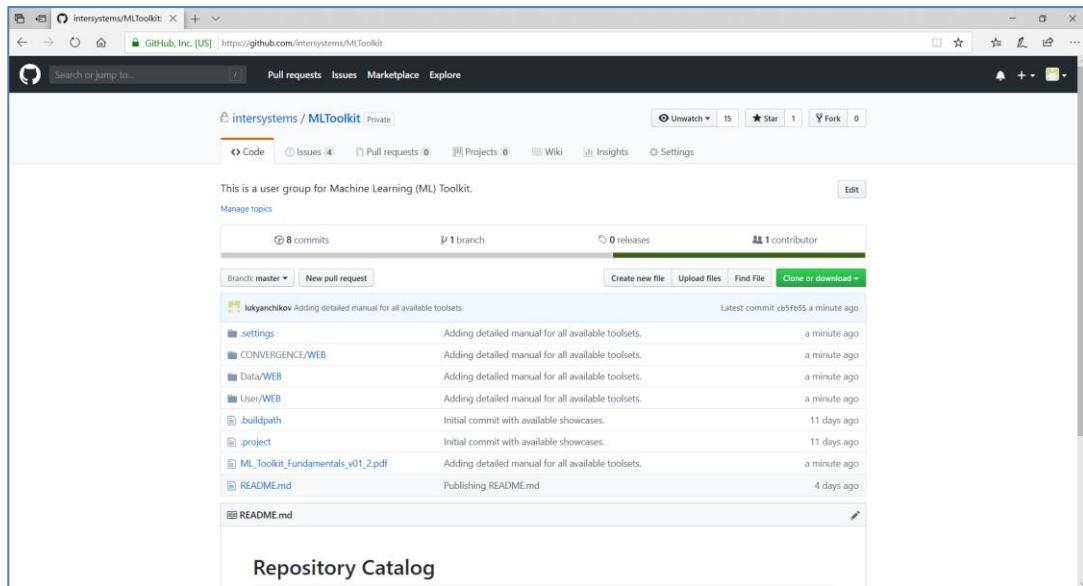


1.2. External Users: ML Toolkit User Group

ML Toolkit user group is a private GitHub repository set up as part of InterSystems corporate GitHub organization. It is addressed to the external users that are installing, learning or are already using ML Toolkit components as specified in this document. To join ML Toolkit user group, please send a short e-mail at the following address:

MLToolkit@intersystems.com and indicate in your e-mail the following details (needed for the group members to get to know and identify you during discussions):

- GitHub account
- Full Name (your first name followed by your last name in Latin script)
- Organization (you are working for, or you study at, or your home office)
- Position (your actual position in your organization, or "Student", or "Independent")
- Country (you are based in)



The screenshot shows the GitHub repository page for 'intersystems/MLToolkit'. The repository has 8 commits, 1 branch, 0 releases, and 1 contributor. The README.md file contains the following content:

```

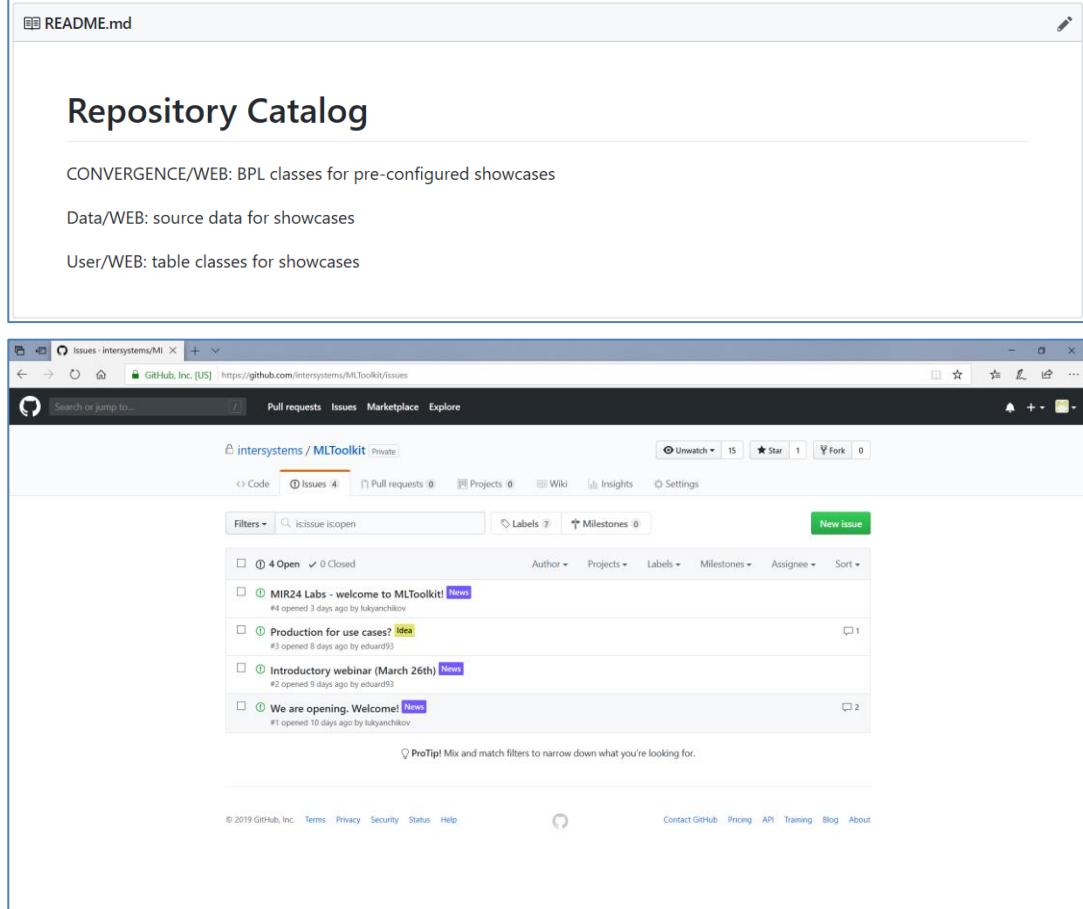
Repository Catalog

CONVERGENCE/WEB: BPL classes for pre-configured showcases

Data/WEB: source data for showcases

User/WEB: table classes for showcases

```

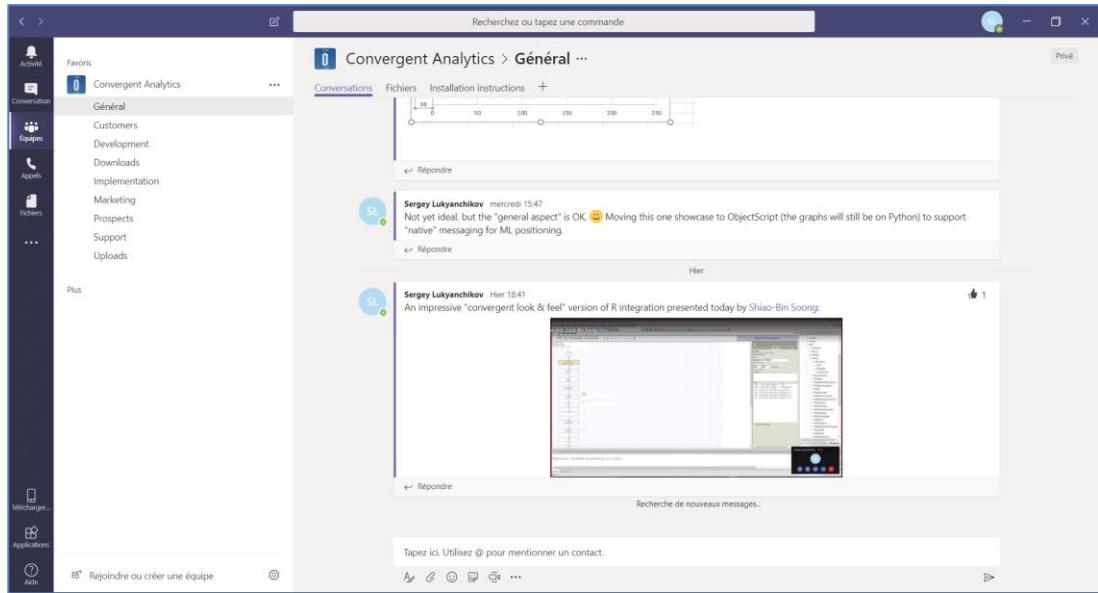


The screenshot shows the GitHub Issues page for the 'intersystems/MLToolkit' repository. There are four open issues listed:

- MIR24 Labs - welcome to MLToolkit! (New)
- Production for use cases? (Idea)
- Introductory webinar (March 26th) (New)
- We are opening. Welcome! (New)

1.3. Internal Users: MS Teams Group

The starting point for accessing all the internal ML Toolkit resources is the MS Teams group Convergent Analytics:



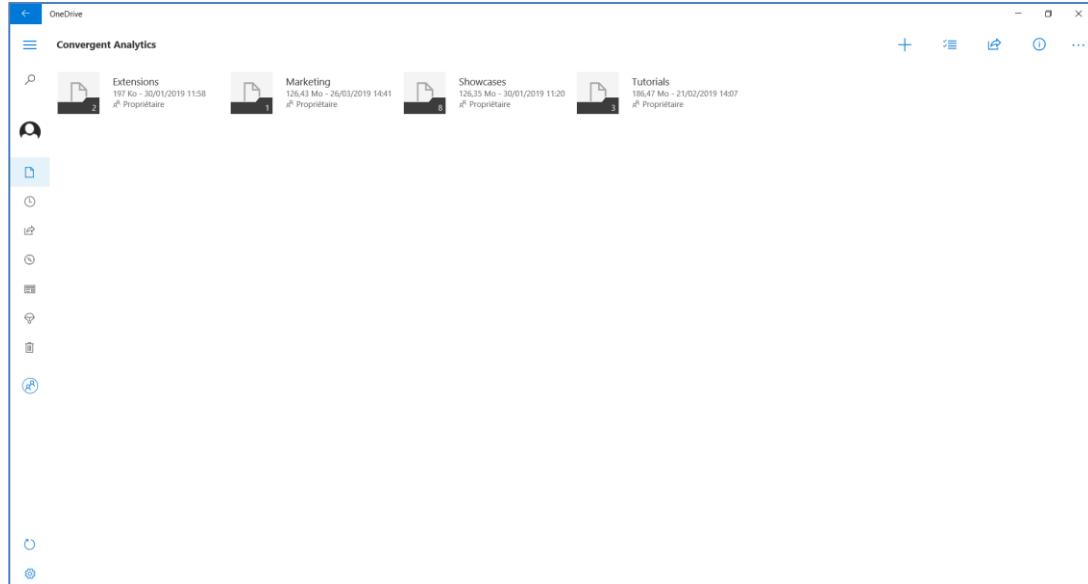
To apply for membership, please use [this link](#).

Once you are in the group, you have access to the following communication channels:

- **General** – receive updates about ML Toolkit feature development and availability, download slide decks and screenshots, learn about webinars and events, dialog with the group
- **Customers** – read and contribute updates about ML Toolkit experience by our valued Customers
- **Development** – stay in sync with vision and thoughts from our code contributors
- **Implementation** – send questions and problems regarding the methodology side of ML Toolkit, leverage help from our colleagues
- **Marketing** – learn about ML Toolkit events, webinars and campaigns driven by our Marketing
- **Downloads** – a shortcut to the copy of the internal OneDrive folder with MS Toolkit extensions, showcases and tutorials
- **Prospects** – share updates on your opportunities that involve ML Toolkit, read updates from our colleagues
- **Support** – send questions and problems regarding the technical side of ML Toolkit, receive answers and support
- **Uploads** – a channel to upload any occasional material relevant to your posts in the other channels

1.4. Internal Users: OneDrive Folder

Having been granted membership in the MS Teams group, you are also granted access to the OneDrive folder Convergent Analytics:



Inside the folder, the following subfolders are maintained:

- **Extensions** – contains functional extensions to IRIS functionality:
 - Python – a set of integration components required to call Python out from IRIS
 - R – a set of integration components required to call R out from IRIS
- **Marketing** – contains marketing materials (recordings, presentations, etc.)
- **Showcases** – contains pre-configured examples and explanatory slides
- **Tutorials** – contains educational materials focused on ML Toolkit

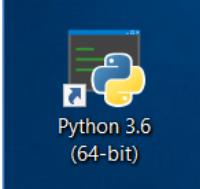
2. Python Tools

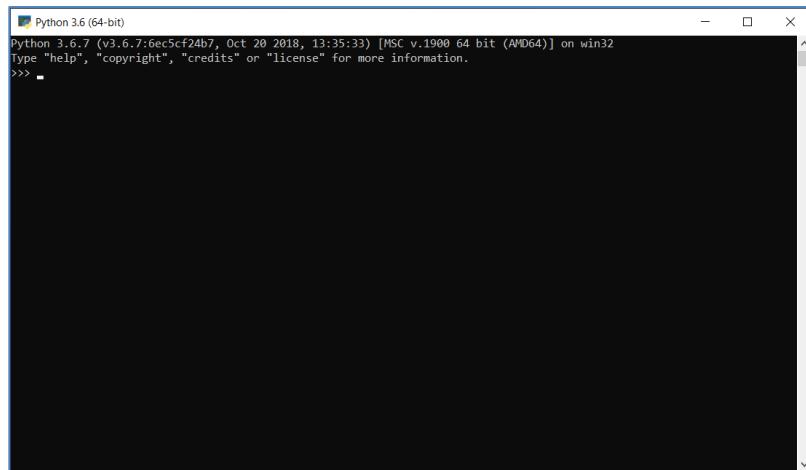
2.1. Python Gateway

Python Gateway for InterSystems data platform. Execute Python code and more from InterSystems IRIS. This toolset brings you the power of Python right into your InterSystems IRIS environment:

- Execute arbitrary Python code
- Seamlessly transfer data from InterSystems IRIS into Python or from Python to InterSystems IRIS
- Build intelligent Interoperability business processes with Python Interoperability Adapter
- Save, examine, modify and restore Python context from InterSystems IRIS

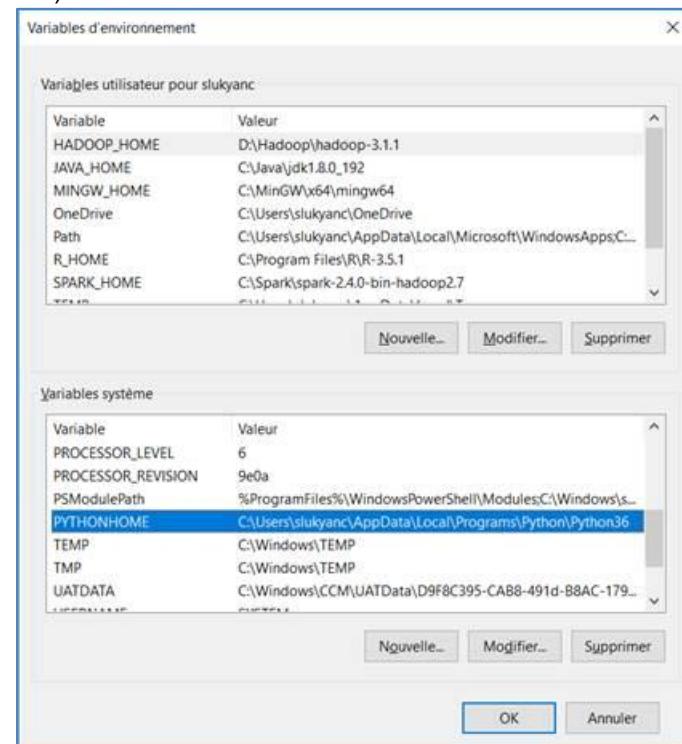
2.2. Installation

1. Install IRIS (**IMPORTANT:** make sure that the IRIS instance service runs under your Windows account and not under a system service account. This is because Python by default installs into your Windows user workspace. You may run IRIS under system service account, but you need to make sure that IRIS has access to Python executable and plugin folders).
2. Install Python and its modules, prepare the environment
 - a. Install Python 3.6.7
 - i. Download Python 3.6.7 64-bit, from the [download page](#). Select the installer that matches your OS and bitness (for example: Windows 64-bit)
 - ii. Install Python 3.6.7 into a default directory (C:\Users\<USER>\AppData\Local\Programs\Python\Python36 on Windows)
 - iii. After the installation, an icon like that appears on your desktop (a Windows-based example):
A blue square icon for Python 3.6 (64-bit). It features the Python logo (a yellow snake) and the text "Python 3.6 (64-bit)" below it.
 - iv. If you double-click it, the following window opens (you can leave it by pressing Ctrl+Z and then pressing Enter):

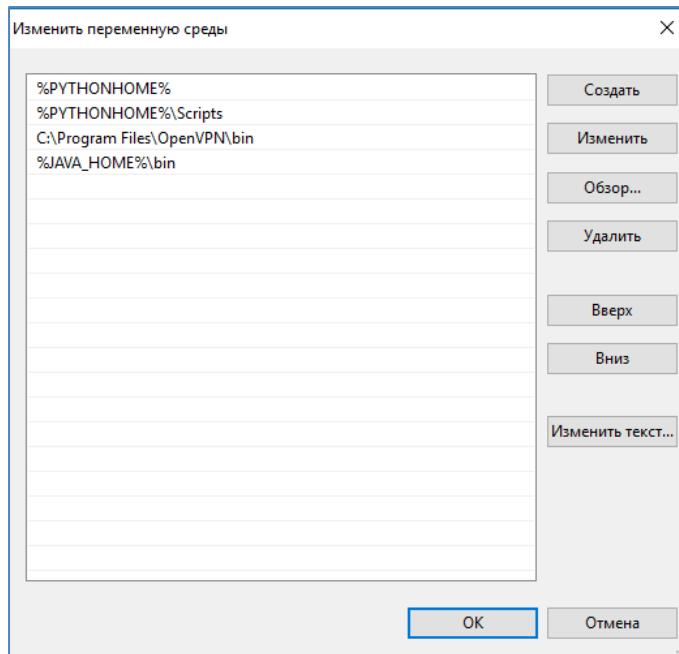


```
Python 3.6.7 (v3.6.7:6ec5cf24b7, Oct 20 2018, 13:35:33) [MSC v.1900 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> 
```

- v. Check that your `PYTHONHOME` system environment variable points to the folder where your Python was installed (for example: `C:\Users\<USER>\AppData\Local\Programs\Python\Python36`):



Also, check that your `PATH` system environment variable includes the path to Python:



If you are on Linux or Mac, check that your `PATH` system environment variable includes `/usr/lib` and `/usr/lib/x86_64-linux-gnu`. Use `/etc/environment` file to set the environment variables.

- vi. Restart your computer for the environment variable changes to take effect.

b. Install Python modules (presuming Python 3.6.7)

- i. Start a command prompt window (e.g., PowerShell or CMD in Windows) and install one by one the following Python modules using `pip install <module name>` command (you need to be connected to the Internet while doing this):

```
Invite de commandes
Microsoft Windows [version 10.0.17763.316]
(c) 2018 Microsoft Corporation. Tous droits réservés.

C:\Users\slukyan>pip install matplotlib
```

- matplotlib
- numpy
- pandas
- seaborn
- sklearn
- statsmodels
- iterools
- tensorflow
- logging

- multiprocessing
- gensim
- dill (required for context harvesting)

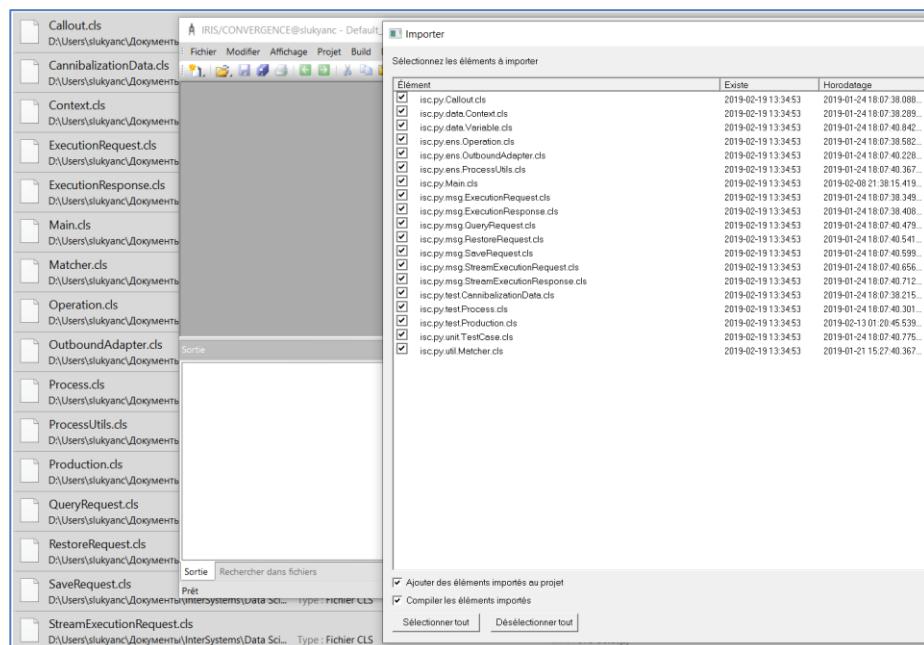
- ii. The warning about pip version not being up to date can be ignored (we can update it any time after the modules installation).
- iii. During the installation various warnings or even error messages can be thrown – we will ignore them for the time being.

3. Import ObjectScript classes using IRIS Studio

- a. in the folder where you keep ML Toolkit installation set, run a file search using *.cls mask:

Callout.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
CannibalizationData.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
Context.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
ExecutionRequest.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
ExecutionResponse.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
Main.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
Matcher.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
Operation.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
OutboundAdapter.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
Process.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
ProcessUtils.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
Production.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
QueryRequest.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
RestoreRequest.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
SaveRequest.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54
StreamExecutionRequest.cls D:\Users\slukyanc\Документы\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 24/01/2019 14:54

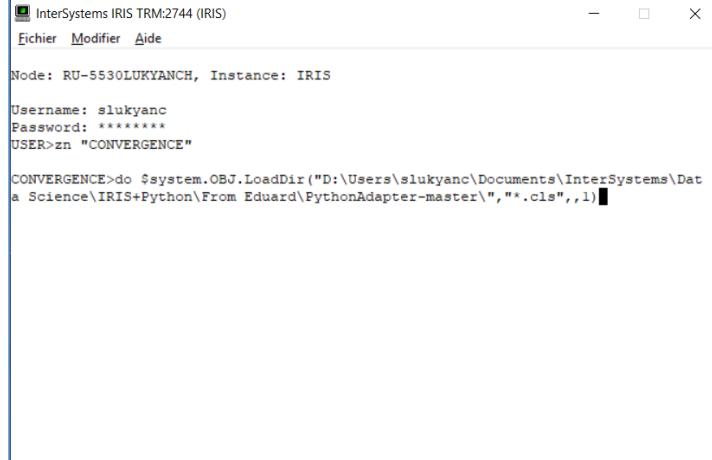
- b. select the files found by the file search above, drag and drop them into your IRIS Studio:



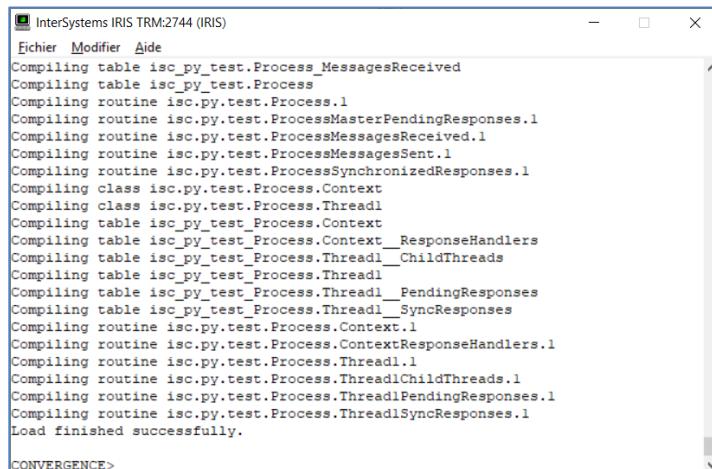
4. Import ObjectScript classes using IRIS Terminal (an alternative to IRIS Studio)

- a. in IRIS Terminal execute the following command (adjust the path to point to the folder where you placed the ML Toolkit class files): do

```
$system.OBJ.LoadDir("C:\path\to\toolkit","*.cls",,1)
```



InterSystems IRIS TRM:2744 (IRIS)
Fichier Modifier Aide
Node: RU-5530LUKYANCH, Instance: IRIS
Username: slukyanc
Password: *****
USER>zn "CONVERGENCE"
CONVERGENCE>do \$system.OBJ.LoadDir("D:\Users\slukyanc\Documents\InterSystems\Dat
a Science\IRIS+Python\From Eduard\PythonAdapter-master\","*.cls",,1)█



InterSystems IRIS TRM:2744 (IRIS)
Fichier Modifier Aide
Compiling table isc_py_test.Process_MessagesReceived
Compiling table isc_py_test.Process
Compiling routine isc_py_test.Process.l
Compiling routine isc_py_test.ProcessMasterPendingResponses.l
Compiling routine isc_py_test.ProcessMessagesReceived.l
Compiling routine isc_py_test.ProcessMessagesSent.l
Compiling routine isc_py_test.ProcessSynchronizedResponses.l
Compiling class isc_py_test.Process.Context
Compiling class isc_py_test.Process.Threadl
Compiling table isc_py_test_Process.Context
Compiling table isc_py_test_Process.Context__ResponseHandlers
Compiling table isc_py_test_Process.Threadl_ChildThreads
Compiling table isc_py_test_Process.Threadl
Compiling table isc_py_test_Process.Threadl_PendingResponses
Compiling table isc_py_test_Process.Threadl_SyncResponses
Compiling routine isc_py_test.Process.Context.l
Compiling routine isc_py.test.Process.ContextResponseHandlers.l
Compiling routine isc_py.test.Process.Threadl.l
Compiling routine isc_py.test.Process.ThreadlChildThreads.l
Compiling routine isc_py.test.Process.ThreadlPendingResponses.l
Compiling routine isc_py.test.Process.ThreadlSyncResponses.l
Load finished successfully.

5. Copy the library file

- a. from the folder where you keep ML Toolkit installation set copy `iscpython.dll` (if you are on Windows), or `iscpython.so` (if you are on Linux), or `iscpython.dylib` (if you are on Mac) to the bin subfolder of your IRIS installation folder, and restart IRIS instance (the library file should be placed into a path returned by `write`

```
##class(isc.py.Callout).GetLib():
```

	Nom	Modifié le	Type	Taille
le	iristrayNLD.dll	23/08/2018 11:54	Extension de l'app...	2 133 Ko
ements	iristrayPTB.dll	23/08/2018 11:54	Extension de l'app...	2 134 Ko
ts	iristrayRUS.dll	23/08/2018 11:54	Extension de l'app...	2 134 Ko
ts	iristrayUKR.dll	23/08/2018 11:54	Extension de l'app...	2 134 Ko
ts	iristrmd.exe	23/08/2018 11:28	Application	15 Ko
ms	IRISQL.dll	23/08/2018 11:53	Extension de l'app...	961 Ko
on	IRISQL64.dll	23/08/2018 10:34	Extension de l'app...	1 080 Ko
ts	IRISUDL.dll	23/08/2018 11:53	Extension de l'app...	434 Ko
ts	IRISUDL64.dll	23/08/2018 10:38	Extension de l'app...	592 Ko
ts	iriswdimj.exe	23/08/2018 10:37	Application	67 Ko
ements	IRISXMLSyn.dll	23/08/2018 11:53	Extension de l'app...	134 Ko
ts	IRISXMLSyn64.dll	23/08/2018 10:32	Extension de l'app...	202 Ko
ts	IRISXSLT.dll	23/08/2018 10:38	Extension de l'app...	216 Ko
ts	ISCEdt32.dll	23/08/2018 11:53	Extension de l'app...	446 Ko
ts	ISCMLink.dll	23/08/2018 11:53	Extension de l'app...	54 Ko
ts	iscpython.dll	24/01/2019 18:09	Extension de l'app...	114 Ko
ts	ISLog.dll	23/08/2018 11:53	Extension de l'app...	106 Ko
ts	itype.dll	23/08/2018 10:51	Extension de l'app...	401 Ko
ts	lcbclient.dll	23/08/2018 11:01	Extension de l'app...	500 Ko
ts	lcbclientnt.dll	23/08/2018 10:56	Extension de l'app...	500 Ko
ts	lcbdotnet.dll	23/08/2018 10:59	Extension de l'app...	136 Ko
ts	lcbdotnett.dll	23/08/2018 11:04	Extension de l'app...	136 Ko
ts	lcbind_msvc120.dll	23/08/2018 11:02	Extension de l'app...	1 332 Ko
ts	lcbindt_msvc120.dll	23/08/2018 10:57	Extension de l'app...	1 333 Ko
ts	lcjni.dll	23/08/2018 11:01	Extension de l'app...	136 Ko
ts	lcjninit.dll	23/08/2018 11:06	Extension de l'app...	136 Ko
ts	ldap.dll	23/08/2018 10:53	Extension de l'app...	29 Ko
ts	libapr-1.dll	23/08/2018 10:56	Extension de l'app...	176 Ko
ts	libeay32.dll	23/08/2018 10:53	Extension de l'app...	2 040 Ko
ts	libhelloworld.dll	18/12/2018 16:32	Extension de l'app...	112 Ko
ts	libhunspell.dll	23/08/2018 10:38	Extension de l'app...	456 Ko
ts	libssh2.dll	23/08/2018 10:53	Extension de l'app...	155 Ko
ts	licmanager.exe	23/08/2018 10:37	Application	87 Ko
ts	mdsdotnet.dll	23/08/2018 11:00	Extension de l'app...	455 Ko
ts	mdsdotnett.dll	23/08/2018 11:05	Extension de l'app...	417 Ko

2.3. Docker

1. To build a Docker image:

- Copy `iscpython.so` into repository root (if it's not there already)
- Execute in repository root `docker build --force-rm --tag intersystemscommunity/irispy:latest`. By default, the image is built upon `intersystems/iris:2019.1.0.510.0-1` image, however you can change that by providing `IMAGE` variable. To build from InterSystems IRIS Community Edition execute: `docker build --build-arg IMAGE=store/intersystems/iris:2019.1.0.510.0-community --force-rm --tag intersystemscommunity/irispy-community:latest`.

2. To run the Docker image execute:

```
docker run -d \
-p 52773:52773 \
-v /<HOST-DIR-WITH-iris.key>/:/mount \
--name irispy \
intersystemscommunity/irispy:latest \
--key /mount/iris.key \
```

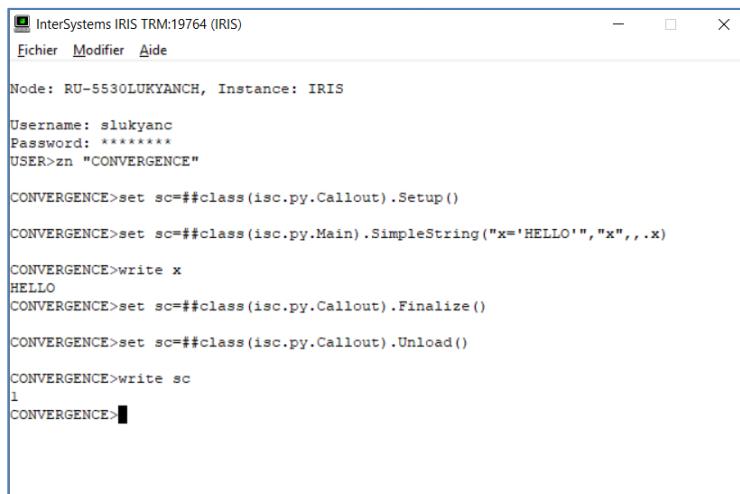
3. Test process `isc.py.test`. Process saves image artifact into `temp` directory. You might want to change that path to a mounted directory. To do that, edit annotation for Correlation Matrix: Graph call, specifying valid filepath for `f.savefig` function.

4. For terminal access execute: `docker exec -it irispy sh`.
5. Access SMP with SuperUser/SYS or Admin/SYS user/password.
6. To stop the container execute: `docker stop irispy && docker rm --force irispy`.

2.4. Use

2.4.1. General Use

1. Execute (once per system start) the following call: set
`sc=#class(isc.py.Callout).Setup()` (add to ZSTART: [docs](#), sample routine available in rtn folder)
2. Continue with calling the main method (can be called multiple times, the context persists): set
`sc=#class(isc.py.Main).SimpleString("x='HELLO'", "x", , .x)`
3. Check the call result: write x
4. To free Python context: set `sc=#class(isc.py.Callout).Finalize()`
5. To free the callout library: set `sc=#class(isc.py.Callout).Unload()`



```

InterSystems IRIS TRM:19764 (IRIS)
Fichier Modifier Aide

Node: RU-5530LUKYANCH, Instance: IRIS

Username: slukyanc
Password: *****
USER>zn "CONVERGENCE"

CONVERGENCE>set sc=#class(isc.py.Callout).Setup()

CONVERGENCE>set sc=#class(isc.py.Main).SimpleString("x='HELLO'", "x", , .x)

CONVERGENCE>write x
HELLO
CONVERGENCE>set sc=#class(isc.py.Callout).Finalize()

CONVERGENCE>set sc=#class(isc.py.Callout).Unload()

CONVERGENCE>write sc
1
CONVERGENCE>

```

6. Overall, `isc.py.Main` is the general interface to Python. It offers the following methods (all return `%Status`):

Code execution (allow execution of arbitrary Python code):

- a. `ImportModule(module, .imported, .alias)` – import a module with an alias
- b. `SimpleString(code, returnVariable, serialization, .result)` – for cases where the code and the variable are both strings
- c. `ExecuteCode(code, variable)` – execute a code (a string or a stream), optionally set the result to a variable
- d. `ExecuteFunction(function, positionalArguments, keywordArguments, variable, serialization, .result)` - execute Python function or method, write result into Python variable, return chosen serialization in result
- e. `ExecuteFunctionArgs(function, variable, serialization, .result, args...)` - execute Python function or method, write result into Python variable, return chosen serialization in result. Builds positionalArguments and keywordArguments and passes them to ExecuteFunction. It is recommended to use ExecuteFunction. More information can be found in [Gateway docs](#).

Data transfer (transfer data into and from Python):

Python -> InterSystems IRIS:

- f. GetVariable(variable, serialization, .stream, useString) – get a serialization of a variable in a stream. If useString is set to 1, and the variable serialization can be fit into a string then the string is returned instead of the stream
- g. GetVariableJson(variable, .stream, useString) – get JSON serialization of a variable
- h. GetVariablePickle(variable, .stream, useString) – get Pickle serialization of a variable

InterSystems IRIS -> Python:

- i. ExecuteQuery(query, variable, type) – create a resultset (of pandas dataframe or list type) from SQL query and set it to a variable
- j. ExecuteGlobal(global, variable, type, start, end, mask, labels, namespace) - transfer global data (from start to end) to Python variable of type: list of tuples or pandas dataframe. For mask and labels arguments specification check class docs and [Data Transfer docs](#)
- k. ExecuteClass(class, variable, type, start, end, properties, namespace) - transfer class data to Python list of tuples or pandas dataframe. properties - comma-separated list of properties to form dataframe from. * and ? wildcards are supported. Defaults to * (all properties). %%CLASSNAME property is ignored. Only stored properties can be used
- l. ExecuteTable(table, variable, type, start, end, properties, namespace) - transfer table data to Python list of tuples or pandas dataframe

ExecuteQuery is universal (any valid SQL query would be transferred into Python). ExecuteGlobal and its wrappers ExecuteClass and ExecuteTable, however, operate with several limitations. But they are much faster (3-5 times faster than ODBC driver and 20 times faster than ExecuteQuery). More information in [Data Transfer docs](#).

Auxiliary (support methods):

- m. GetVariableInfo(variable, serialization, .defined, .type, .length) – get information on a variable: is it defined, type and serialization length
- n. GetVariableDefined(variable, .defined) - is variable defined
- o. GetVariableType(variable, .type) - get variable FQCN
- p. GetStatus() – returns the last occurred exception in Python and clears it
- q. GetModuleInfo(module, .imported, .alias) – get the module alias and its currently imported status
- r. GetFunctionInfo(function, .defined, .type, .docs, .signature, .arguments) - get function information

7. Possible serialization functions:

- a. ##class(isc.py.Callout).SerializationStr – a serialization by str() function if set to 1 (the default value is 0)
- b. ##class(isc.py.Callout).SerializationRepr – a serialization by repr() function if set to 1 (the default value is 1)

2.4.2. Shell

To open Python shell: do `##class(isc.py.util.Shell).Shell()`. To exit press enter.

2.4.3. Context Persistence

Python context can be persisted into IRIS and restored later. There are the following functions:

1. Save the context: set
`sc=##class(isc.py.data.Context).SaveContext(.context,maxLength,mask,verbose)` where maxLength is the maximum length of the saved variable. If the variable serialization is longer than that, it will be ignored. Set to 0 to get them all. The other parameters: mask is a comma-separated list of the variables to save (special symbols * and ? are recognized), verbose specifies displaying the context after saving and context is the resulting Python context. Get the context ID with `context.%Id()`
2. Display the context: do
`##class(isc.py.data.Context).DisplayContext(id)` where id is the ID of a stored context. Leave empty to display the current context
3. Restore the context: do
`##class(isc.py.data.Context).RestoreContext(id,verbose,clear)` where clear kills the currently loaded context if set to 1

A context is saved into `isc.py.data` package and can be viewed/edited by SQL and object methods.

2.4.4. IRIS Interoperability Adapter

IRIS Interoperability adapter `isc.py.ens.Operation` enables interaction with a Python process from Interoperability productions. The following requests are currently supported:

1. Execute Python code via `isc.py.msg.ExecutionRequest` and get the response via `isc.py.msg.ExecutionResponse` with requested variable values as strings
2. Execute Python code via `isc.py.msg.StreamExecutionRequest` and get the response via `isc.py.msg.StreamExecutionResponse` with requested variable values as streams
3. Transfer data into Python from an SQL query with `isc.py.msg.QueryRequest` and get the response via `Ens.Response`
4. Save a Python context with `isc.py.msg.SaveRequest` and get the response via `Ens.StringResponse` with the context ID
5. Restore a Python context with `isc.py.msg.RestoreRequest`

Check request/response classes documentation for more details.

Settings:

- Initializer - select a class implementing `isc.py.init.Abstract`. It can be used to load functions, modules, classes and so on. It would be executed at process start
- PythonLib - (Linux only) if you see loading errors set it to `libpython3.6m.so` or even to a full path to the shared library

Note: `isc.py.util.BPEmulator` class is added to allow easy testing of Python Interoperability business processes. It can execute business process (python parts) in a current job.

Variable substitution

All business processes inheriting from `isc.py.ens.ProcessUtils` can use `GetAnnotation(name)` method to get the value of activity annotation by activity name. Activity annotation can contain variables which would be calculated on ObjectScript side before being passed to Python. This is the syntax for variable substitution:

- `#{class:method:arg1:...:argN}` - execute method
- `#{expr}` - execute ObjectScript code

Check test `isc.py.test.Process` business process for example in Correlation Matrix: Graph activity:

```
f.savefig(r'#{process.WorkDirectory}SHOWCASE${%PopulateUtils:Integer:1:100}.png')
```

In this example:

- `#{process.WorkDirectory}` returns `WorkDirectory` property of process object which is an instance of `isc.py.test.Process` class and current business process
- `#{%PopulateUtils:Integer:1:100}` calls `Integer` method of `%PopulateUtils` class passing arguments 1 and 100, returning random integer in range 1...100

2.4.5. Test Business Process

To use the sample business process and production:

1. In OS shell execute: `pip install pandas matplotlib seaborn`
2. Execute this code in terminal to populate the data: do
`##class(isc.py.test.CannibalizationData).Import()`
3. In the sample business process `isc.py.test.Process`, edit the annotation for the Correlation Matrix: Graph call specifying a valid file path in `f.savefig` function
4. Save and compile the business process
5. Start `isc.py.test.Production` production
6. Send an empty `Ens.Request` message to `isc.py.test.Process`

Notes:

- If you want to use ODBC connectivity – on Windows install `pyodbc`: `pip install pyodbc`, on Linux install `apt-get install unixodbc unixodbc-dev python-pyodbc`
- If you want to use JDBC connectivity – install `JayDeBeAPI`: `pip install JayDeBeApi`, on Linux you may need to install system packages beforehand: `apt-get install python-apt`
- If you are getting errors similar to `undefined symbol: _Py_TrueStruct`, in `isc.py.ens.Operation` operation set `PythonLib` setting to `libpython3.6m.so` or even to a full path to the shared library. Check troubleshooting section for more details.
- In the sample business process `isc.py.test.Process` edit the annotations for ODBC or JDBC connection calls specifying a correct connection string
- In production, for the sample business process `isc.py.test.Process` set `ConnectionType` setting to a preferred connection type (defaults to `RAW`, change only if you need to test xDBC connectivity)

2.4.6. Unit Tests

To run unit tests, execute:

```
set repo=##class(%SourceControl.Git.Utils).TempFolder()
set
^UnitTestRoot=##class(%File).SubDirectoryName(##class(%File).SubDirectoryName(##class(%File).SubDirectoryName(repo,"isc"),"py"),"unit",1)
set sc=##class(%UnitTest.Manager).RunTest(,/nodelete")
```

2.4.7. ZPY Command

Install [this ZLANG routine](#) to add zpy command:

```
zpy "import random"
zpy "x=random.random()"
zpy "x"
>0.4157151243124494
```

Argumentless zpy command opens Python shell.

2.4.8. Limitations

There are several limitations:

1. Module reinitialization. Some modules may only be loaded once per process lifetime (i.e. numpy). While Finalization clears the context of the process, repeated loading of such libraries terminates the process. Discussions: [1](#), [2](#).
2. Variables. Do not use these variables: zzz*. Please report any leakage of zzz* variables. System code should always clear them.
3. Functions. Do not redefine these functions: zzz*()
4. Context persistence. Only pickled/dill variables can be restored correctly. Module imports are supported.

2.5. Development

Development of ObjectScript code is done via [cache-tort-git](#) in UDL mode.

Development of C code is done in Eclipse.

2.5.1. Commits

Commits should follow the pattern: module: description issue. List of modules:

- Callout - C and ObjectScript callout interface in isc.py.Callout.
- API - terminal API, mainly isc.py.Main.
- Gateway - proxy classes generation.
- Proxyless Gateway - isc.py.gw.DynamicObject class.
- Interoperability - support utilities for Interoperability Business Processes.
- Tests - unit tests and test production.
- Docker - containers.
- Docs - documentation.

2.5.2. Building

[Windows:](#)

1. Install [MinGW-w64](#) – you will need make and gcc
2. In mingw64\bin directory, rename mingw32-make.exe to make.exe
3. Set GLOBALS_HOME environment variable to the root of IRIS installation
4. Set PYTHONHOME environment variable to the root Python installation (usually C:\Users\<User>\AppData\Local\Programs\Python\Python3<x>)
5. Open MinGW shell (mingw64env.cmd)
6. In <Repository>\c\ execute make

Linux:

It is recommended to use Linux OS that uses Python 3.X by default, i.e. Ubuntu 18.04.1 LTS. Skip steps 1 and probably 2 if your OS has Python 3.6X as default (to check Python version: python3 -version or python -version or python3.6 -version)

1. Add Python 3.6 repo: add-apt-repository ppa:jonathonf/python-3.6 and apt-get update
2. Install: apt install python3.6 python3.6-dev libpython3.6-dev build-essential
3. Set GLOBALS_HOME environment variable to the root of IRIS installation
4. Set environment variable PYTHONVER to the Python version you want to build, i.e. export PYTHONVER=3.6
5. In <Repository>\c\ execute make

Mac OS X:

1. Install Python 3.6 and gcc compiler
2. Set GLOBALS_HOME environment variable to the root of IRIS installation
3. Set environment variable PYTHONVER to the Python version you want to build, i.e. export PYTHONVER=3.6
4. In <Repository>\c\ execute:


```
gcc -Wall -Wextra -fpic -O3 -fno-strict-aliasing -Wno-unused-parameter -I/Library/Frameworks/Python.framework/Versions/${PYTHONVER}/Headers -I${GLOBALS_HOME}/dev/iris-callin/include -c -o iscpython.o iscpython.c
      gcc -dynamiclib -L/Library/Frameworks/Python.framework/Versions/${PYTHONVER}/lib -L/usr/lib -lpython${PYTHONVER}m -lpthread -ldl -lutil -lm -Xlinker iscpython.o -o iscpython.dylib
```

If you have a Mac please update makefile so we can build Mac version via make

2.5.3. Troubleshooting

1. <DYNAMIC LIBRARY LOAD> exception
 - a. Check that the OS has the correct Python installed:


```
import sys
          sys.version
```

 The result should contain Python 3.6.7 and 64-bit. If it does not, [install Python 3.6.7 64-bit](#)
 - b. Check OS-specific installation steps. Make sure that the path relevant for IRIS (usually, the system PATH) contains Python installation directory

- c. Make sure that IRIS can access Python installation
2. Module not found error

Sometimes you may get module not found error. This is how you can fix it. Each step constitutes a complete solution requiring IRIS restart and check on whether the problem has been solved

- a. Check that OS and IRIS use the same Python. Open both Pythons, execute the below script in each and verify that the versions are the same:

```
import sys
ver=sys.version
ver
```

If they are not the same, search for the Python executable that is used by IRIS

- b. Check that the module is, in fact, installed. Open OS shell, execute `python` (or `python3` or `python36` in Linux) and in the open shell execute `import <module>`. If it fails with an error, in OS shell run `pip install <module>`. Note that a module name for `import` and a module name for `pip` can be different
- c. If you are sure that the module is installed, compare the paths used by Python (nothing to do with system PATH). Get the path with:

```
import sys
path=sys.path
path
```

The returned paths should be the same. If they are not the same, read how `PYTHONPATH` (Python) is formed here and adjust your OS environment to form it correctly, i.e. set `PYTHONPATH` (system environment variable) to `C:\Users\<USER>\AppData\Roaming\Python\Python36\site-packages` or to other directories where your modules reside (plus other missing directories)

- d. Compare Python paths again, and if they are still not the same or the problem persists, add the missing paths explicitly to `isc.py.ens.OutboundAdaptor` init code (for Interoperability) and on process start (for Callout wrapper):

```
do ##class(isc.py.Main).SimpleString("import sys")
do
##class(isc.py.Main).SimpleString("sys.path.append('C:\\\\U
sers\\\\<USER>\\\\AppData\\\\Roaming\\\\Python\\\\Python36\\\\site-
packages')")
```

3. undefined symbol: `_Py_TrueStruct` or similar errors

- a. Check `ldconfig` and adjust it to point to the directory with Python shared library
- b. If it fails:
 - i. for Interoperability: in `isc.py.ens.Operation` operation set `PythonLib` setting at `libpython3.6m.so` or even at a full path to the shared library
 - ii. for Callout wrapper: on process start call do

```
##class(isc.py.Callout).Initialize("libpython3.6m.so"), alternatively pass a full path to the shared library
```

4. PyODBC on Linux and Mac

- a. Install unixodbc: `apt-get install unixodbc-dev`
 - b. Install PyODBC: `pip install pyodbc`. To install on Linux, read the [notes](#)
 - c. Set the connection string: `cnxn=pyodbc.connect('Driver=/<IRIS directory>/bin/libirisodbcu35.so;Server=localhost;Port=51773;database=USER;UID=_SYSTEM;PWD=SYS'),autocommit=True)`
- Call `set sc =`
`#class(isc.py.util.Installer).ConfigureTestProcess(user,`
`pass, host, port, namespace)` to configure test process automatically.

2.6. Gateway Functionality

2.6.1. Execute function

Executes function by name. This API consists of two methods:

- `ExecuteFunction`
- `ExecuteFunctionArgs`

The difference between them is caller signature. `ExecuteFunction` accepts `%List`, `%Collection`, `AbstractArray` and JSON object separated into positional and keyword arguments. `ExecuteFunctionArgs` accepts `args...` and parses them into positional and keyword arguments. After that `ExecuteFunctionArgs` calls `ExecuteFunction`.

It is caller responsibility to escape argument values. Use `isc.py.util.Converter` class to escape:

- string
- boolean
- date
- time
- timestamp

ExecuteFunction:

`ExecuteFunction` method from `isc.py.Main` class. Signature:

- function - name of the function to invoke. Can be nested, i.e.
`random.randint`
- variable - name of the python variable to write the result to.
- positionalArguments - positional arguments for Python function. Can be one of:
 - `$lb(val1, val2, ..., valN)`
 - `%Collection`, `AbstractIterator` object
 - JSON array
- keywordArguments - keyword arguments for Python function. Can be one of:
 - `$lb($lb(name1, val1), $lb(name2, val2), ..., $lb(nameN, valN))`
 - `%Collection`, `AbstractArray` object
 - flat JSON object
- serialization - how to serialize the result

- **result** - write the result into this variable

All arguments besides **function** are optional.

Here is an example of how it works:

```

set sc = ##class(isc.py.Main).ImportModule("random", ,.random)

set posList = $lb(1, 100)
set posCollection = ##class(%ListOfDataTypes).%New()
do posCollection.Insert(1)
do posCollection.Insert(100)
set posDynamic = [1, 100]

for positionalArguments = posList,posCollection,posDynamic {
    set sc = ##class(isc.py.Main).ExecuteFunction(random _
".randint", positionalArguments,,.result)
    write result,
}

set kwList = $lb($lb("a", 1), $lb("b", 100))
set kwCollection = ##class(%ArrayOfDataTypes).%New()
do kwCollection.SetAt(1, "a")
do kwCollection.SetAt(100, "b")
set kwDynamic = { "a": 1, "b": 100}

for kwArguments = kwList,kwCollection,kwDynamic {
    set sc = ##class(isc.py.Main).ExecuteFunction(random _
".randint", ,kwArguments,,.result)
    write result,
}

set posList = $lb(1)
set kwDynamic = {"b": 100}
set sc = ##class(isc.py.Main).ExecuteFunction(random _ ".randint",
posList, kwDynamic,,.result)
write result,

set posList =
##class(isc.py.util.Converter).EscapeStringList($lb("Positional:
{0} {1}! Keyword: {name}, {name2}", "Hello", "World"))
set kwDynamic =
{"name":(##class(isc.py.util.Converter).EscapeString("Alice")),
"name2":(##class(isc.py.util.Converter).EscapeString("Bob"))}
set sc = ##class(isc.py.Main).ExecuteFunction("str.format",
posList, kwDynamic,,.result)
write result,

```

ExecuteFunctionArgs:

ExecuteFunctionArgs method from `isc.py.Main` class. Signature:

function - name of the function to invoke. Can be nested, i.e. `random.randint`

variable - name of the python variable to write the result to.

serialization - how to serialize the result

result - write the result into this variable

args... - function arguments.

ExecuteFunctionArgs attempts to determine correct positional and keyword arguments from the function signature (if available). It is recommended to call ExecuteFunction directly if ExecuteFunctionArgs is unable to construct a correct argument spec (and opens an issue). Example:

```
set sc = ##class(isc.py.Main).ImportModule("random", ,.random)
set sc = ##class(isc.py.Main).ExecuteFunctionArgs(random _
".randint", , ,.result, 1, 100)
write result,!  
  

set string =
##class(isc.py.util.Converter).EscapeString("Positional: {0}, {1},
{2}, {3}")
set arg1 = ##class(isc.py.util.Converter).EscapeString("Hello")
set arg2 = ##class(isc.py.util.Converter).EscapeString("World")
set arg3 = ##class(isc.py.util.Converter).EscapeString("Alice")
set arg4 = ##class(isc.py.util.Converter).EscapeString("Bob")
set sc =
##class(isc.py.Main).ExecuteFunctionArgs("str.format",,,.result,
string, arg1, arg2, arg3, arg4)
write result,!  
  

set string =
##class(isc.py.util.Converter).EscapeString("Positional: {0} {1}!
Keyword: {name}, {name2}")
set arg1 = ##class(isc.py.util.Converter).EscapeString("Hello")
set arg2 = ##class(isc.py.util.Converter).EscapeString("World")
set kwargs = "*** _ {"name":"Alice","name2":"Bob"}.%ToJSON()
set sc = ##class(isc.py.Main).ExecuteFunctionArgs("str.format",,,.
result, string, arg1, arg2, kwargs)
write result,!  

```

2.6.2. Proxyless Gateway

Proxyless gateway allows user to bind Python variables to InterSystems IRIS variables.

This allows user to:

- Get/Set object properties
- Call object methods
- Serialize variable to: Str, Repr, Pickle, Dill, JSON, Dynamic Object.

Example.

1. Load Python class Person: do ##class(isc.py.init.Test).Initialize(,1)

Note: here is Person class definition for reference:

```
class Person(object):
    def __init__(self, name, age, city):
        self.name = name
        self.age = age
        self.city = city
    def getAge(self):
        return self.age
    def getAgePlus(self, add):
        return self.age + add
```

2. Create Proxy variable:

```
set obj = ##class(isc.py.gw.DynamicObject).%New("Person", "p1",
"'Ed'", "25", "'Test'")
```

In this call we create Python variable p1 of Person class and pass three methods to constructor 'Ed', 25 and 'Test'.

3. Now we can interact with the object, let us get and set some properties:

```
write obj.name
set obj.name="Bob"
write obj.name
write obj.age
```

4. We can set some new properties too (unlike ExecuteFunction values are escaped automatically if %EscapeOnSet property is 1, which is default. You can also set properties to other dynamic objects. In that case the unescaped Python variable name would be used):

```
set obj.pet = "Dog"
write obj.pet
```

5. And we can call object methods:

```
write obj.getAge()
write obj.getAgePlus(10)
```

6. Finally we can convert an object:

```
set sc = obj.%ToJSON(.json)
set sc = obj.%ToDynObj(.dynObj)
set sc = obj.%ToPickle(.pickle)
set sc = obj.%ToStream(.stream)
```

To create a proxy object from an existing proxy object just skip the type argument:

```
kill obj
set p1 = ##class(isc.py.gw.DynamicObject).%New(, "p1")
```

Module objects can be proxied this way too:

```
set module = "random"
set sc = ##class(isc.py.Main).ImportModule(module)
set random = ##class(isc.py.gw.DynamicObject).%New(,module)
write random.randint(1,100)
```

Now for a more complex example. In case of primitives (int, bool, str, float) a proxy object returns a serialized value. Otherwise (if a method call or a variable get returns a complex type) it returns another proxy object pointing to that result.

```

set sc = ##class(isc.py.Main).ImportModule("numpy",, "np")
set np = ##class(isc.py.gw.DynamicObject).%New(, "np")
set arr = ##class(isc.py.gw.DynamicObject).%New("np.array",
"arr", "[[1.5,2],[4,5]]")
set exp = np.exp(arr)
write $replace(exp.%GetString(), $c(10), $c(13,10))

And here is an example of setting a property to a proxy object:
do ##class(isc.py.init.Test).Initialize(,1)
set obj = ##class(isc.py.gw.DynamicObject).%New("Person", "p1",
"'Ed'", "25", "'Test'")
set obj2 = ##class(isc.py.gw.DynamicObject).%New("Person", "p2",
"'Bob'", "22", "'Test2'")
write obj.%GetJSON()

set obj.relative = obj2
set obj3 = obj.relative
write obj3.%GetJSON()

```

You can use `%EscapeOnSet` and `%EscapeOnCall` properties and `%IsPrimitive` method to affect default serialization behavior.

2.7. Data Transfer

Transfer data into and from Python. All the methods are defined in `isc.py.Main`. All methods return `%Status`.

2.7.1. Python -> InterSystems IRIS

- `GetVariable(variable, serialization, .stream, useString)` - get serialization of variable in stream. If `useString` is 1 and variable serialization can fit into string, then string is returned instead of the stream
- `GetVariableJson(variable, .stream, useString)` - get JSON serialization of variable
- `GetVariablePickle(variable, .stream, useString, useDill)` - get Pickle (or Dill) serialization of variable

2.7.2. InterSystems IRIS -> Python

Load data from InterSystems IRIS to Python. All these methods support data transfer from any local namespace. `isc.py` package must be available in your working namespace.

ExecuteQuery

`ExecuteQuery(query, variable, type, namespace)` - transfer results from any valid SQL query into Python. It is the slowest method of data transfer. Use it if `ExecuteGlobal` and its wrappers are unavailable.

Arguments:

- `query` - sql query
- `variable` - target variable on a Python side

- `type` - list or Pandas dataframe

ExecuteGlobal

`ExecuteGlobal(global, variable, type, start, end, mask, labels, namespace)` - transfer global data to Python.

Arguments:

- `global` - global name without ^
- `variable` - target variable on a Python side
- `type` - list or Pandas dataframe
- `start` - initial global key. Must be integer.
- `end` - final global key. Must be integer.
- `mask` - string, mask for global values. Mask may be shorter than the number of global value fields (in this case fields at the end would be skipped). How to format mask:
 - + use field as is
 - – skip field
 - b - boolean (0 - False, anything else - True)
 - d - date (from \$horolog, on Windows only from 1970, on Linux from 1900 see notes for details)
 - t - time (\$horolog, seconds since midnight)
 - m - (moment) timestamp string in YEAR-MONTH-DAY HOUR:MINUTE:SECOND format.
- `labels` - %List of column names, first element is key column name. Therefore: list length must be mask symbol length + 1

ExecuteClass

Wrapper for `ExecuteGlobal`. Effectively it parses compiled class definition, constructs `ExecuteGlobal` arguments and calls it.

`ExecuteClass(class, variable, type, start, end, properties, namespace)` - transfer class data to Python list of tuples or pandas dataframe.

`properties` - comma-separated list of properties to form dataframe from. * and ? wildcards are supported. Defaults to * (all properties). %%CLASSNAME property is ignored. Only stored properties can be used.

Arguments:

- `class` - class name
- `variable` - target variable on a Python side
- `type` - list or Pandas dataframe
- `start` - initial object id. Must be integer.
- `end` - final object id. Must be integer.
- `properties` - comma-separated list of properties to form dataframe from. * and ? wildcards are supported. Defaults to * (all properties). %%CLASSNAME property is ignored. Only stored properties can be used.

All properties transferred as is except properties of %Date, %Time, %Boolean and %TimeStamp types. They are converted to respective Python datatypes.

ExecuteTable

Wrapper for `ExecuteClass`. Translates table name to class name and calls `ExecuteClass`. Signature:

`ExecuteTable(table, variable, type, start, end, properties, namespace)` - transfer table data to Python list of tuples or pandas dataframe.

Arguments:

- `table` - table name.

Other arguments are passed as is to `ExecuteClass`.

Notes

- `ExecuteGlobal`, `ExecuteClass` and `ExecuteQuery` generally offer the same speed (as the time to parse class definition is negligible)
- `ExecuteGlobal` is 3-5 times faster than ODBC driver and up to 20 times faster than `ExecuteQuery` on measurable workloads (>0.01 second)
- `ExecuteGlobal`, `ExecuteClass` and `ExecuteQuery` only work on the globals with this structure: `^global(key) = $lb(prop1, prop2, ..., propN)` where key must be an integer
- For `ExecuteGlobal`, `ExecuteClass` and `ExecuteQuery` supported %Date range equals mktime range ([Windows](#): 1970-01-01, [Linux](#): 1900-01-01, [Mac](#)). Use `%TimeStamp` to transfer dates outside of this range
- For `ExecuteGlobal`, `ExecuteClass` and `ExecuteQuery` all arguments besides source (global, class, table) and variable are optional

Examples

Let's say we have `isc.py.test.Person` class. Here's how we can use all methods of data transfer:

```
// All the ways to transfer data
set global = "isc.py.test.PersonD"
set class = "isc.py.test.Person"
set table = "isc_py_test.Person"
set query = "SELECT * FROM isc_py_test.Person"

// Common arguments
set variable = "df"
set type = "dataframe"
set start = 1
set end = $g(^isc.py.test.PersonD, start)

// Approach 0: ExecuteGlobal without arguments
set sc = ##class(isc.py.Main).ExecuteGlobal(global, variable _ 0,
type)

// Approach 1: ExecuteGlobal with arguments
// For global transfer labels are not calculated automatically
// globalKey - is global subscript
set labels = $lb("globalKey", "Name", "DOB", "TS", "RandomTime",
"AgeYears", "AgeDecimal", "AgeDouble", "Bool")
```

```
// mask is 1 element shorter than labels because "globalKey" is
// global subscript label
// Here we want to skip %%CLASSNAME field
set mask = "-+dmt+++b"

set sc = ##class(isc.py.Main).ExecuteGlobal(global, variable _ 1,
type, start, end, mask, labels)

// Approach 2: ExecuteClass
set sc = ##class(isc.py.Main).ExecuteClass(class, variable _ 2,
type, start, end)

// Approach 3: ExecuteTable
set sc = ##class(isc.py.Main).ExecuteTable(table, variable _ 3,
type, start, end)

// Approach 4: ExecuteTable
set sc = ##class(isc.py.Main).ExecuteQuery(query, variable _ 4,
type)

You can call this method: do ##class(isc.py.test.Person).Test() to check how
these data transfer methods work.
```

3. R Tools

3.1. R Gateway

R Gateway for InterSystems data platform. Execute R code and more from InterSystems IRIS. This toolset brings you the power of R right into your InterSystems IRIS environment:

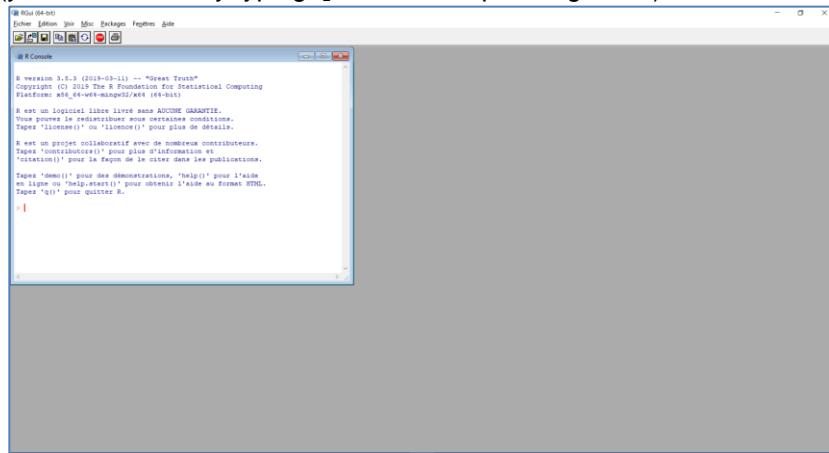
- Execute arbitrary R code
- Seamlessly transfer data from InterSystems IRIS into R or from R to InterSystems IRIS
- Build intelligent Interoperability business processes with R Interoperability Adapter

3.2. Installation

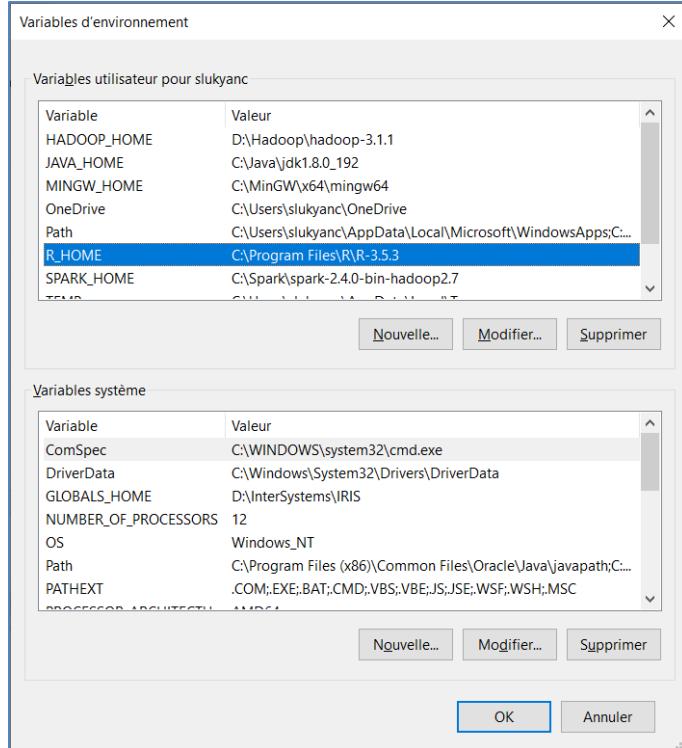
1. Install IRIS.
2. Install R and its modules, prepare the environment
 - a. Install R
 - i. Download R, from the [download page](#). Select the installer that matches your OS (for example: Windows). By default, Windows installer will install both 32 and 64-bit executables/dlls – we recommend accepting the default installation.
 - ii. Install R into a default directory (`C:\Program Files\R\R-#.#.#` on Windows)
 - iii. After the installation, two icons like the ones below appear on your desktop (a Windows-based example):



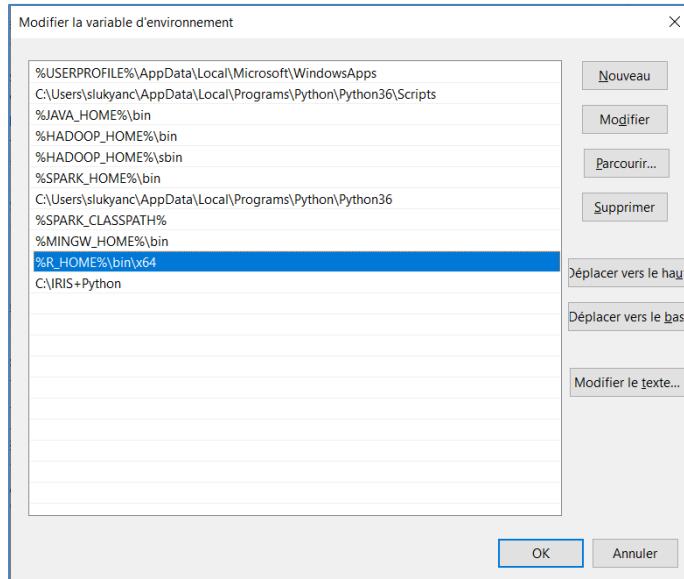
- iv. If you double-click the one that corresponds to 64 bits, the following window opens (you can leave it by typing `q()` and then pressing Enter):



- v. Check that your `R_HOME` system environment variable points to the folder where your R was installed (for example: `C:\Program Files\R\R-#.#.#`):



Also, check that your PATH system environment variable includes the path to R:

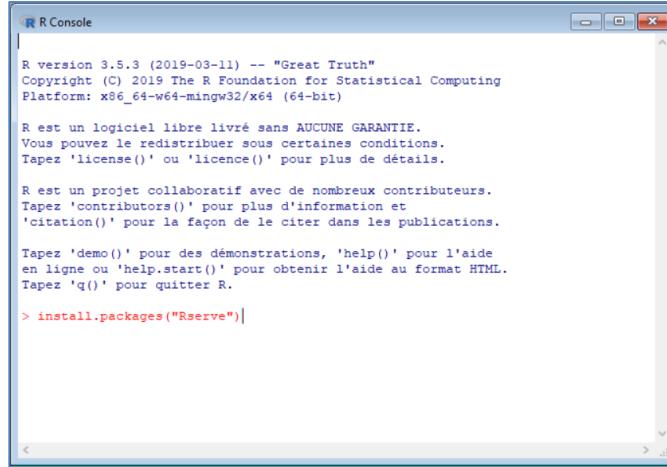


If you are on Linux or Mac, check that your PATH system environment variable includes /usr/lib and /usr/lib/x86_64-linux-gnu. Use /etc/environment file to set the environment variables.

vi. Restart your computer for the environment variable changes to take effect.

b. Install R packages

- i. Start R as administrator (right-click on the R icon in Windows and select running it as administrator) and install one by one the following R packages using `install.packages ("<module name>")` command (you need to be connected to the Internet while doing this):



```
R version 3.5.3 (2019-03-11) -- "Great Truth"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R est un logiciel libre livré sans AUCUNE GARANTIE.
Vous pouvez le redistribuer sous certaines conditions.
Tapez 'license()' ou 'licence()' pour plus de détails.

R est un projet collaboratif avec de nombreux contributeurs.
Tapez 'contributors()' pour plus d'information et
'citation()' pour la façon de le citer dans les publications.

Tapez 'demo()' pour des démonstrations, 'help()' pour l'aide
en ligne ou 'help.start()' pour obtenir l'aide au format HTML.
Tapez 'q()' pour quitter R.

> install.packages("Rserve")|
```

- Rserve
- igraph

ii. Before proceeding to a module download, R will prompt you to a download server selection list.

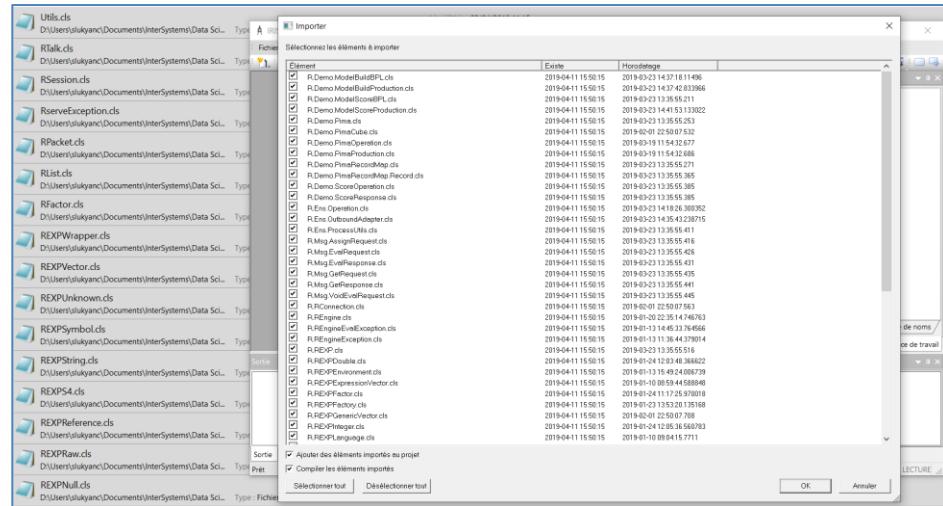
iii. After the download server has been selected, the module download starts.

3. Import ObjectScript classes using IRIS Studio

- in the folder where you keep ML Toolkit installation set, run a file search using *.cls mask:

	Utils.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 891 octet(s)
	RTalk.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 13,0 Ko
	RSession.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 1,21 Ko
	RserveException.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 2,17 Ko
	RPacket.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 1,10 Ko
	RList.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 4,13 Ko
	RFactor.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 5,00 Ko
	REXPWrapper.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 57 octet(s)
	REXPVector.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 1003 octet(s)
	REXPUnknown.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 415 octet(s)
	REXPsymbol.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 807 octet(s)
	REXPString.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 1,56 Ko
	REPS4.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 105 octet(s)
	REXPReference.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 2,64 Ko
	REXPRaw.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 1,03 Ko
	REXPNull.cls D:\Users\slukyanc\Documents\InterSystems\Data Sci...	Type : Fichier CLS	Modifié le : 02/04/2019 11:15 Taille : 349 octet(s)

- select the files found by the file search above, drag and drop them into your IRIS Studio:



4. Import ObjectScript classes using IRIS Terminal (an alternative to IRIS Studio)
- in IRIS Terminal execute the following command (adjust the path to point to the folder where you placed the ML Toolkit class files): do
- ```
$system.OBJ.LoadDir("C:\path\to\toolkit","*.cls",,1)
```

```
InterSystems IRIS TRM:4620 (IRIS)
Fichier Modifier Aide
Node: RU-5530LUKYANCH, Instance: IRIS
Username: slukyanc
Password: *****
USER>zn "CONVERGENCE"
CONVERGENCE>do $system.OBJ.LoadDir("D:\Users\slukyanc\Documents\InterSystems\Dat a Science\IRIS+R\From Shiao-Bin","*.cls",,1)
```

```
InterSystems IRIS TRM:4620 (IRIS)
Fichier Modifier Aide
Compiling routine R.Demo.PimaCube.StarTricepsSkinRg867535915.1
Compiling routine R.Demo.PimaCube.StarSerumInsulinRg4081394933.1
Compiling routine R.Demo.ModelBuildBPL.Context.1
Compiling routine R.Demo.ModelBuildBPL.ContextResponseHandlers.1
Compiling routine R.Demo.ModelScoreBPL.Context.1
Compiling routine R.Demo.ModelScoreBPL.ContextResponseHandlers.1
Compiling routine R.Demo.PimaCube.Fact.1
Compiling routine R.Test.Process.Context.1
Compiling routine R.Test.Process.ContextResponseHandlers.1
Compiling routine R.Demo.ModelBuildBPL.Thread1.1
Compiling routine R.Demo.ModelBuildBPL.Thread1ChildThreads.1
Compiling routine R.Demo.ModelBuildBPL.Thread1PendingResponses.1
Compiling routine R.Demo.ModelBuildBPL.Thread1SyncResponses.1
Compiling routine R.Demo.ModelScoreBPL.Thread1.1
Compiling routine R.Demo.ModelScoreBPL.Thread1ChildThreads.1
Compiling routine R.Demo.ModelScoreBPL.Thread1PendingResponses.1
Compiling routine R.Demo.ModelScoreBPL.Thread1SyncResponses.1
Compiling routine R.Test.Process.Thread1.1
Compiling routine R.Test.Process.Thread1ChildThreads.1
Compiling routine R.Test.Process.Thread1PendingResponses.1
Compiling routine R.Test.Process.Thread1SyncResponses.1
Load finished successfully.
```

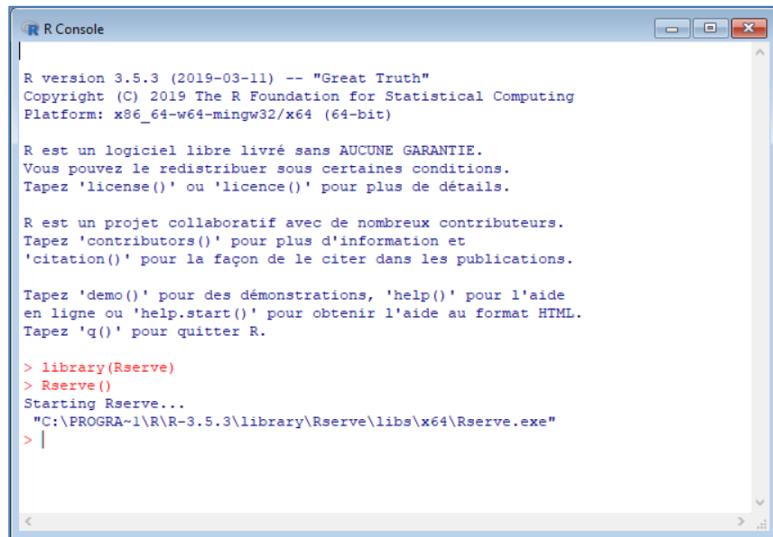
### 3.3. Use

#### 3.3.1. General Use

- Before starting the use of R Gateway, start the Rserve component by typing in R:  

```
library(Rserve)
```

```
Rserve()
```



```
R version 3.5.3 (2019-03-11) -- "Great Truth"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R est un logiciel libre livré sans AUCUNE GARANTIE.
Vous pouvez le redistribuer sous certaines conditions.
Tapez 'license()' ou 'licence()' pour plus de détails.

R est un projet collaboratif avec de nombreux contributeurs.
Tapez 'contributors()' pour plus d'information et
'citation()' pour la façon de le citer dans les publications.

Tapez 'demo()' pour des démonstrations, 'help()' pour l'aide
en ligne ou 'help.start()' pour obtenir l'aide au format HTML.
Tapez 'q()' pour quitter R.

> library(Rserve)
> Rserve()
Starting Rserve...
"C:\PROGRA~1\R\R-3.5.3\library\Rserve\libs\x64\Rserve.exe"
> |
```

2. Class `R.Ens.Operation` is the general interface to R. It offers the following methods (all return `%Status`):
  - a. `Assign(request, .response)` – assigns values to a named variable in R
  - b. `Eval(request, .response)` – execute an R code and evaluate in IRIS-mappable terms code's named variables
  - c. `VoidEval(request, .response)` – execute an R code without evaluating in IRIS-mappable terms code's named variables
  - d. `Get(request, .response)` – retrieves values from a named variable in R
  - e. `AssignSQL(request, .response)` – execute an SQL statement locally and assign the resultset as dataframe to a named variable in R

### 3.3.2. IRIS Interoperability Adapter

IRIS Interoperability adapter `R.Ens.Operation` enables interaction with an R process from Interoperability productions. The following requests are currently supported:

1. Assign values to a named variable in R via `R.Msg.AssignRequest`
2. Execute an R code and evaluate in IRIS-mappable terms code's named variables via `R.Msg.EvalRequest` and get the response via `R.Msg.EvalResponse`
3. Execute an R code without evaluating in IRIS-mappable terms code's named variables with `R.Msg.VoidEvalRequest`
4. Retrieve values from a named variable in R with `R.Msg.GetRequest` and get the response via `R.Msg.GetResponse`
5. Execute an SQL statement and assign its resultset as dataframe to a named variable in R via `R.Msg.AssignSQLRequest`

The above requests contain a serial object `R.Msg.RData`. This class maps all common R data types to IRIS Interoperability requests so that R data elements can show up in message traces. It also provides some convenient methods to work with data stored in globals. The details are documented in class `R.Msg.RData`.

Check respective request/response classes documentation for more details.

### 3.3.3. Notes

- If you want to use ODBC connection, on Windows install package RODBC or packages DBI and odbc
- If you want to use JDBC connection, install RJDBC

### 3.3.4. Sample Business Process and Production

There are two sample productions in the demo package R.Demo:  
`ModelBuildProduction` is used to train, test and save a model which predicts the likelihood of a patient being diagnosed with diabetes. `ModelScoreProduction` shows how to make predictions for new patients using the previously saved model.

To build a predictive model:

1. Run the following command to populate the sample dataset:

```
do ##class(R.Demo.Pima).Import()
```

2. Configure and start production `ModelBuildProduction`.

3. Send an empty Ens.Request message to `ModelBuildBPL` process.

To make new predictions:

1. Configure and start production `ModelScoreProduction`.

2. Copy sample file `pima-new-dataset.csv` to the file path of `Enslib.RecordMap.Service.FileService`.

3. If the likelihood of a patient being diagnosed with diabetes is more than 50%, an alert will be sent.

Note that on Windows machines, each instance of `Rserve` can only support one R Gateway connection. Be sure to stop the current running production before start a new one.

### 3.3.5. Unit Tests

All test code is under `R.Test` package. `R.Test.API` contains the tests for the underlying API. `R.Test.Production` and `R.Test.Process` cover the usage of `R.Ens.Operation` and `R.Msg` messages.

### 3.3.6. Limitations

1. It is known that each instance of `Rserve` on Windows only supports one R Gateway connection. To support multiple connections, you can start multiple instance of `Rserve`, each with a different port.
2. `Rserve` is thread-safe across connections, but `*eval*` methods are not thread-safe within one connection. This means that multiple threads should not use the same connection unless they guarantee that no `*eval*` calls are run in parallel.

### 3.3.7. Rserve Configuration

The default `Rserve` server port is 6311. It can be started with a different port through a configuration file. On Linux or Mac computer, the configuration file is at `/etc/Rserve.conf`. On Windows machine, it can be specified by `--RS-conf` option during start up. For example, `Rserve (args="--RS-conf C:\\\\InterSystems\\\\Rserve.conf")`.

All supported configuration directives are listed below:

- `log.io` (boolean) if enabled the content of I/O messages is logged in debug mode (has no effect in normal mode)
- `daemon` (boolean) if enabled `Rserve` will daemonize
- `close.all.stdio` (boolean) if enabled, `stdout/stderr` are closed and redirected to `/dev/null`
- `msg.id` (boolean) if enabled `Rserve` uses message ID in the headers
- `remote` (boolean) if enabled servers listen on all external interfaces so they can be used over the network, otherwise just loopback is bound
- `tag.argv` (boolean) if enabled the program name is changed after forking to distinguish the server instance from working sessions
- `forward.stdio` (boolean) if enabled `stdout/stderr` is forwarded using OOB send
- `ulog` string, defining destination for logging. Can be either a path to a local socket or either of `tcp://<host>[:port]` or `udp://<host>:[port]` to send logging to an external server (if port is not specified 514 is used). The protocol is compatible with `syslogd`
- `keep.alive` (boolean) if enabled `SO_KEEPALIVE` is enabled on the socket
- `switch.qap.tls` (boolean) if enabled `TLS` switch is allowed for QAP connections
- `qap.oc` (boolean, alternate name `rserve.oc`) if enabled Object-Capability mode of the QAP protocol is used
- `console.oob` (boolean) if enabled R console I/O generates OOB send messages (OOB must be enabled for this to have effect)
- `console.input` (boolean) if enabled R `ReadConsole` API invokes OOB message to query the client for input (has no effect unless `console.oob` is also enabled)
- `websockets.qap.oc` (boolean) if enabled the WebSockets QAP server uses Object-Capability mode
- `random.uid` (boolean) if enabled random `uid` is used when switching `uid` after connect
- `random.gid` (boolean) if enabled random `gid` is used when switching `gid` after connect
- `random.uid.range` (string of the form `xxx-yyy`) range for randomly generated `uid`
- `auto.uid` (boolean) if enabled `uid` is determined from the password file on authentication
- `auto.gid` (boolean) if enabled `gid` is determined from the password file on authentication
- `default.uid` (integer) `uid` to use as fallback if it cannot be determined during authentication
- `default.gid` (integer) `gid` to use as fallback if it cannot be determined during authentication
- `oob.idle.interval` (integer) interval in seconds after which an "idle" OOB send packet is sent to the client (mostly to prevent proxies from dropping the connection)
- `qap.port` (integer, alternative name `port`) port to use by the QAP server
- `ipv6` (boolean) if enabled servers listen on IPv6
- `http.upgrade.websockets` (boolean) if enabled HTTP server allows upgrade to the WebSockets protocol on the same connection

- `http.raw.body` (boolean) if enabled HTTP callback send raw (unparsed) body
- `websockets.port` (integer) port to use for the WebSockets server
- `http.port` (integer) port to use for the HTTP server
- `tls.key` (path) path to the file containing the RSA key to be used for TLS
- `tls.ca` (path) path to the file (in PEM format) containing Certificate Authority certificates to be registered with TLS
- `tls.cert` (path) path to the file (in PEM format) containing the certificate to use for TLS servers
- `pid.file` (path) path to the file that will hold the PID of the `Rserve` server process once started
- `rsa.key` (path) path to the RSA key to use for RSA authentication
- `qap.tls.port` (integer, alternative name `tls.port`) port of the secure server running QAP wrapped in TLS
- `http.tls.port` (integer, alternative name `https.port`) port of the secure server running HTTP wrapped in TLS
- `websockets.tls.port` (integer) port of the secure server running WebSockets wrapped in TLS
- `qap` (boolean, alternative name `rserve`) if enabled QAP server is started
- `websockets.qap` (boolean) if enabled WebSockets mode with QAP protocol is started
- `websockets.text` (boolean) if enabled WebSockets mode with text protocol is started (deprecated and discouraged)
- `websockets` (boolean) if enabled all WebSocket modes are enabled
- `maxinbuf` (integer) limit for the size of incoming packets in kB
- `source` (path) path to a file to use via `source()` prior to starting the servers
- `eval` (string) the string will be parsed and evaluated in the global environment prior to starting the servers
- `maxsendbuf` (integer) limit for the size of the output buffer in kB
- `su` (string, one of `now`, `server` or `client`) determines the time at which a user switch is performed
- `http.user` (string) username to switch to when running the HTTP server
- `https.user` (string) username to switch to when running the HTTP/TLS server
- `websockets.user` (string) username to switch to when running the WebSockets server
- `uid` (integer/oct) user ID to switch to
- `gid` (integer/oct) group ID to switch to
- `chroot` (path) path to use as a chroot jail
- `umask` (integer/oct) umask to use
- `allow` (string) IP address to add to the white-list (only IPv4 is supported at this point)
- `control` (boolean) if enabled control commands (`server.eval`, `server.source`, `server.shutdown`) are allowed

- `shutdown` (boolean) if enabled `CMD_shutdown` is allowed (it does not affect shutdown via `SIGINT`)
- `workdir` (path) path to the working directory used as root for per-connection directories
- `workdir.clean` (boolean) if enabled the working directory for a closed connection is removed upon exit even if it is dirty
- `workdir.mode` (integer/oct) mode (UNIX permissions) for the per-connection working directories
- `workdir.parent.mode` (integer/oct) mode for the root of all per-connection working directories
- `encoding` (string) string encoding to use in the protocol. It must be an encoding recognized by R `iconv` facilities
- `socket` (path) path to the local socket used for QAP
- `sockmod` (integer/oct) mode for the local socket
- `pwdfile` (path) path to the password file. The file is expected to contain one user/password pair per line, separated by a whitespace
- `auth.function` (string) name of an R function to use for authentication instead of the built-in `Rserve` facilities
- `auth` (boolean or require - only first character is checked) if enabled authentication is required before any other command can be used
- `interactive` (boolean) if enabled R is run in interactive mode, otherwise not
- `plaintext` (boolean) if enabled plain-text authentication is allowed
- `oob` (boolean) if enabled out-of-band (OOB) messages (send and msg) can be used by the R code run in the session
- `fileio` (boolean) if enabled file I/O QAP commands are allowed
- `r.control` (boolean, alternative name `r-control`) if enabled the R session can use self-command to issue control commands (as opposed to control commands initiated by the client)
- `cachepwd` (boolean or indefinitely) controls caching of passwords. If enabled the passwords file is read on connection (before `su` is executed), if disabled it is read at the time of authentication and if indefinitely then it only read at the time of server start

### 3.3.8. Development

Development of ObjectScript code can be done with IRIS Studio, or Eclipse. R scripts development and test are done with R Studio and R Terminal.

### 3.3.9. Troubleshooting

1. Takes a long time to make an R Gateway connection or it simply hangs while connecting: a connection may be still active on the specified port. Just kill the `Rserve` process associated with the port.

### 3.3.10. Gateway Functionality

The public interfaces of R Gateway are all under class `R.RConnection`.

`R.Ens.OutboundAdapter` and `R.Ens.Operation` encapsulate the API details in the context of Ensemble production.

- To create an R Gateway connection:

```
set c = ##class(R.RConnection).%New(host, port)
```

- To assign a double value to a R variable:

```
set x = ##class(R.REXPDouble).%New(3.0)
do c.assign("x", x)
```

- To evaluate R script:

```
do c.eval("y<-sqrt(x)")
```

- To retrieve the value of an R variable:

```
set y = c.get("y")
```

The API do not return a status code. It is advised to wrap all ObjectScript code in a try/catch block. Please refer to `R.RConnection` class documentation for details. More examples can be found in `R.Test.API`.

`R.REXP` is the super class of all supported R data types. It contains many useful utility methods:

- `getJSON()` converts the R data type to an IRIS `%DynamicObject`
- `toString()` is the shortened string representation of the R data type
- `toDebugString()` is the string representation of the underlying `%DynamicObject`
- `asDoubleMatrix()` converts the internal matrix data to an IRIS multidimensional array
- `createDoubleMatrix()` takes a multidimensional array and converts it to an R matrix object
- `createDataFrame()` creates an internal R data frame object
- `createDataFrameFromSQL()` executes the SQL statement and converts the result set to an R data frame object

There are more generic utility methods in `R.Utils` class.

To simulate an R terminal from an IRIS terminal: `do`

```
##class(R.Utils).RTerminal().
```

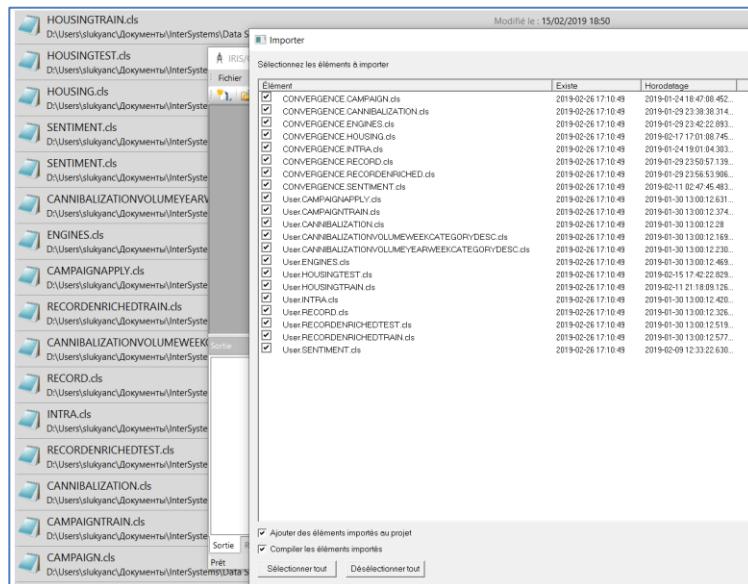
## 4. IRIS Interoperability Tools

### 4.1. Installation

1. Import ObjectScript classes using IRIS Studio (**IMPORTANT:** run installation for all the tool dependencies that are assumed by the content you are installing in this section)
  - a. in the folder where you keep ML Toolkit showcases, run a file search using \*.cls mask:

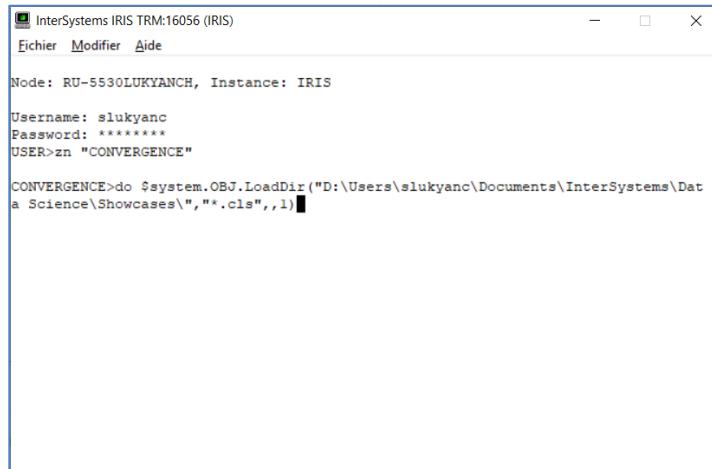
|                                                                                                                                                                                                           |                                                   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|
|  CANNIBALIZATION.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                           | Modifié le : 26/02/2019 16:23<br>Taille : 4,89 Ko |
|  HOUSINGTRAIN.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                              | Modifié le : 26/02/2019 16:17<br>Taille : 4,37 Ko |
|  HOUSINGTEST.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                               | Modifié le : 15/02/2019 18:50<br>Taille : 10,9 Ko |
|  HOUSING.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                                   | Modifié le : 15/02/2019 18:50<br>Taille : 11,0 Ko |
|  SENTIMENT.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                                 | Modifié le : 11/02/2019 02:48<br>Taille : 11,2 Ko |
|  SENTIMENT.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                                 | Modifié le : 09/02/2019 12:32<br>Taille : 2,81 Ko |
|  CANNIBALIZATIONVOLUMEYEARWEEKCATEGORYDESC.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS | Modifié le : 30/01/2019 12:20<br>Taille : 4,48 Ko |
|  ENGINES.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                                   | Modifié le : 30/01/2019 12:20<br>Taille : 4,66 Ko |
|  CAMPAIGNAPPLY.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                             | Modifié le : 30/01/2019 12:20<br>Taille : 3,51 Ko |
|  RECORDENRICHEDTRAIN.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                       | Modifié le : 30/01/2019 12:20<br>Taille : 10,7 Ko |
|  CANNIBALIZATIONVOLUMEWEEKCATEGORYDESC.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS     | Modifié le : 30/01/2019 12:20<br>Taille : 4,46 Ko |
|  RECORD.cls<br>D:\Users\slukyan\Documents\InterSystems\Data Sci... Type : Fichier CLS                                    | Modifié le : 30/01/2019 12:20<br>Taille : 3,82 Ko |
|  INTRA.cls<br>D:\Users\slukyan\Documents\InterSystems>Data Sci... Type : Fichier CLS                                    | Modifié le : 30/01/2019 12:20<br>Taille : 3,94 Ko |
|  RECORDENRICHEDTEST.cls<br>D:\Users\slukyan\Documents\InterSystems>Data Sci... Type : Fichier CLS                      | Modifié le : 30/01/2019 12:20<br>Taille : 10,7 Ko |
|  CANNIBALIZATION.cls<br>D:\Users\slukyan\Documents\InterSystems>Data Sci... Type : Fichier CLS                         | Modifié le : 30/01/2019 12:20<br>Taille : 3,12 Ko |

- b. select the files found by the file search above, drag and drop them into your IRIS Studio:



2. Import ObjectScript classes using IRIS Terminal (an alternative to IRIS Studio)

- a. in IRIS Terminal execute the following command (adjust the path to point to the folder where you placed the ML Toolkit showcase class files): do  
`$system.OBJ.LoadDir("C:\path\to\showcases","*.cls",,1)`



```
InterSystems IRIS TRM:16056 (IRIS)
Fichier Modifier Aide

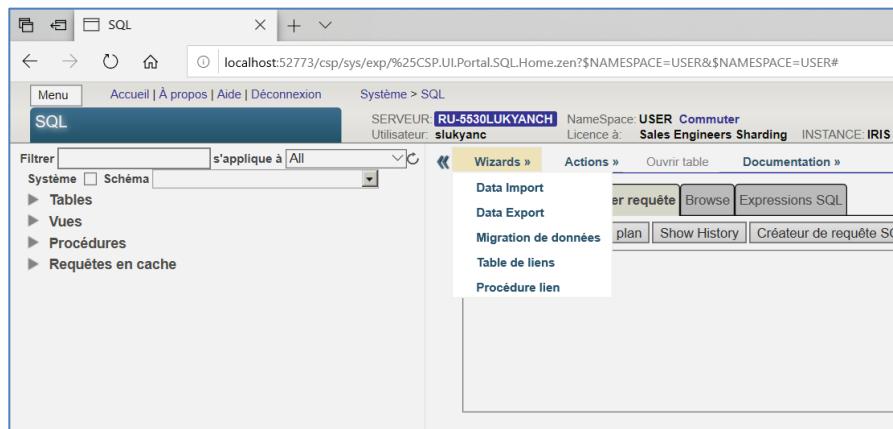
Node: RU-5530LUKYANCH, Instance: IRIS

Username: slukyanc
Password: *****
USER>zn "CONVERGENCE"

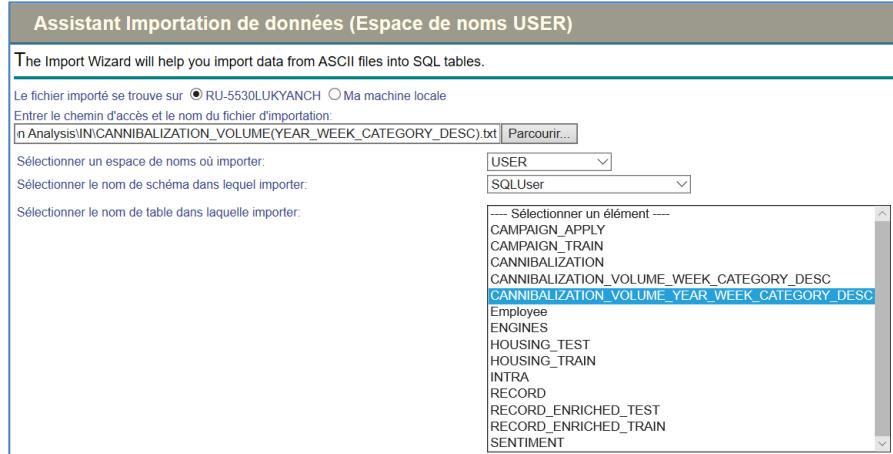
CONVERGENCE>do $system.OBJ.LoadDir("D:\Users\slukyanc\Documents\InterSystems\Dat
a Science>Showcases","*.cls",,1)
```

### 3. Import showcase data

- b. from the folder where you keep ML Toolkit showcase input data, import the showcase dataset into its respective IRIS table using Data Import wizard:



The screenshot shows the SQL interface with the 'Wizards' tab selected. Under 'Data Import', the 'Importation de données' option is highlighted. Below it, there are buttons for 'Plan', 'Show History', and 'Créateur de requête SQL'.

This dialog box provides instructions for importing data from ASCII files into SQL tables. It includes fields for selecting the import file (set to 'RU-5530LUKYANCH'), choosing the target schema ('USER'), and selecting the target table ('CANNIBALIZATION\_VOLUME\_YEAR\_WEEK\_CATEGORY\_DESC'). A scrollable list at the bottom shows other available tables like 'Employee', 'ENGINES', and 'HOUSING\_TEST'.

**Assistant Importation de données (Espace de noms USER)**

Quelles colonnes voulez-vous inclure depuis SQLUser.CANNIBALIZATION\_VOLUME\_YEAR\_WEEK\_CATEGORY\_DESC?

| Disponible                               | Sélectionné                              |
|------------------------------------------|------------------------------------------|
| ----- Sélectionner un ou plusieurs ----- | ----- Sélectionner un ou plusieurs ----- |
| YEAR                                     | YEAR                                     |
| WEEK                                     | WEEK                                     |
| BRATWURST                                | BRATWURST                                |
| CERVELAS                                 | CERVELAS                                 |
| CHARCUTERIE                              | CHARCUTERIE                              |
| DAUERFLEISCHWAREN                        | DAUERFLEISCHWAREN                        |
| GEFLUEGEL                                | GEFLUEGEL                                |
| GERAEUCHERTES_ZUM_KOCHEN                 | GERAEUCHERTES_ZUM_KOCHEN                 |
| HACKFLEISCH                              | HACKFLEISCH                              |

**Assistant Importation d'... (Espace de noms USER)**

Selectionnez les options décrivant comment les données sont stockées dans le fichier ASCII pour SQLUser.CANNIBALIZATION\_VOLUME\_YEAR\_WEEK\_CATEGORY\_DESC.

Délimiteur séparant les colonnes?  Onglet  Espace  L'onglet file  Caractère [ ]

La première ligne contient-elle les en-têtes de colonne?

Chaine entre guillemets: [ ]

Format date: MM/DD/YYYY

Format heure: hh:mm:ss

Format horodatage: ODBC Format

Désactiver validation?

Reportez la création d'index à l'aide de %SortBegin%/%SortEnd%

Aperçu données

| Colonne | NAME                     | Type            | Largeur* |
|---------|--------------------------|-----------------|----------|
| 1       | YEAR                     | %Library.String | 255      |
| 2       | WEEK                     | %Library.String | 255      |
| 3       | BRATWURST                | %Library.String | 320      |
| 4       | CERVELAS                 | %Library.String | 320      |
| 5       | CHARCUTERIE              | %Library.String | 320      |
| 6       | DAUERFLEISCHWAREN        | %Library.String | 320      |
| 7       | GEFLUEGEL                | %Library.String | 320      |
| 8       | GERAEUCHERTES_ZUM_KOCHEN | %Library.String | 320      |
| 9       | HACKFLEISCH              | %Library.String | 320      |
| 10      | INNEREIJEN_DIVERSES      | %Library.String | 320      |
| 11      | KALB                     | %Library.String | 320      |
| 12      | KANINCHEN                | %Library.String | 320      |
| 13      | LAMM                     | %Library.String | 320      |
| 14      | MARINADEN                | %Library.String | 320      |

< Précédent Suivant > Terminer Annuler

**Assistant Importation de données (Espace de noms USER)**

Vérifiez vos sélections ci-dessous. Ensuite, cliquez sur le bouton Terminer.

Nom fichier: D:\Users\slukyanci\Documents\InterSystems\Datasets\Showcases\004\_Retail\_Cannibalization\_Analysis\SQL\CANNIBALIZATION\_VOLUME(YEAR\_WEEK\_CATEGORY\_DESC).txt

Jeu de caractères: <Périphérique par défaut>

Schéma: SQLUser

Table: CANNIBALIZATION\_VOLUME\_YEAR\_WEEK\_CATEGORY\_DESC

Colonnes délimitées par: Onglet

La première ligne contient-elle les en-têtes de colonne? Oui

Chaine entre guillemets: double

Format date: MM/DD/YY/YY

Format heure: hh:mm:ss

Format horodatage: ODBC Format

Désactiver validation? Non

Reportez la création d'index à l'aide de %SortBegin%/%SortEnd% Non

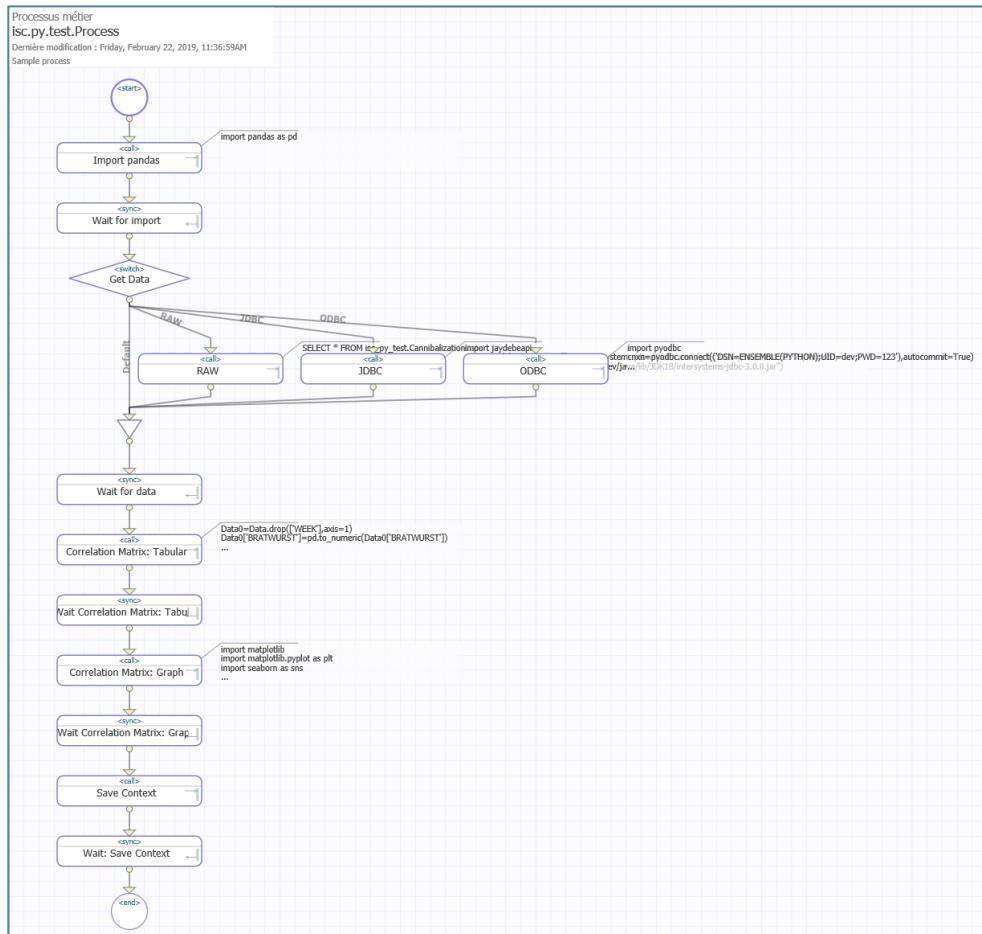
Finish the wizard and the showcase dataset will be imported.

## 4.2. Use

### 4.2.1. General Use

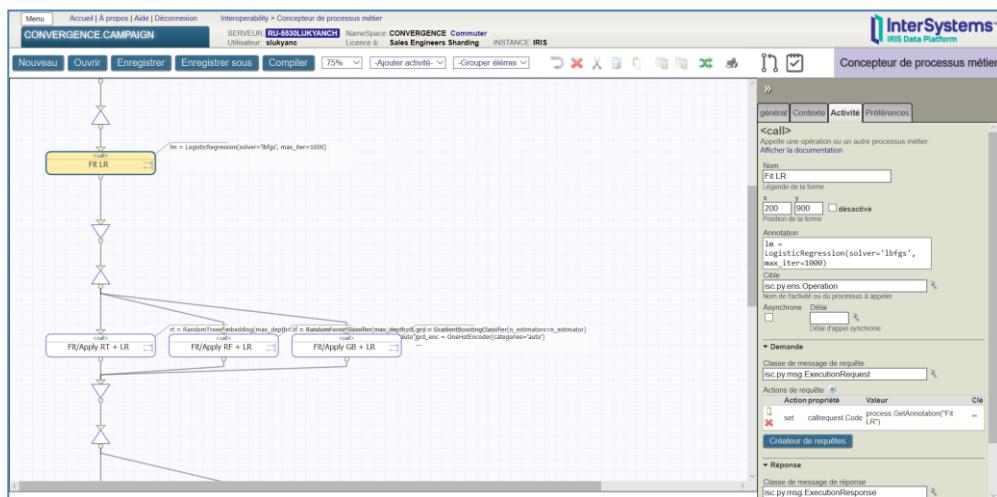
The visual business process designer in IRIS Interoperability is one of the tools for implementing adaptive analytical processes (find more at the following [link](#)). All toolsets (e.g., Python Tools, R Tools, etc.) from ML Toolkit can be leveraged in an adaptive analytical process, one process combining tools coming from various toolsets and generating analytical objects (models, matrices, vectors, datasets, graphs, etc.) in the respective tool's context. As such, extracting analytical objects from the external tools' contexts to IRIS Interoperability context and saving them in IRIS as global variables ("globals"), objects and tables, creates a vast space for analytical interoperability and integration.

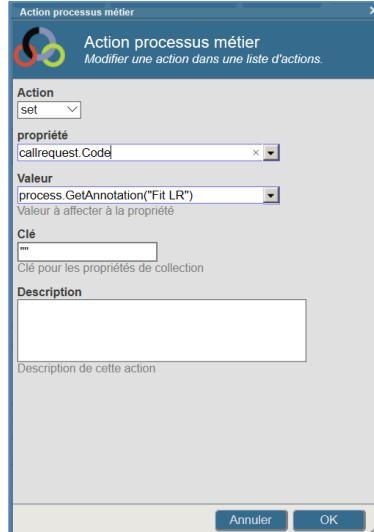
## 4.2.2. Python Gateway



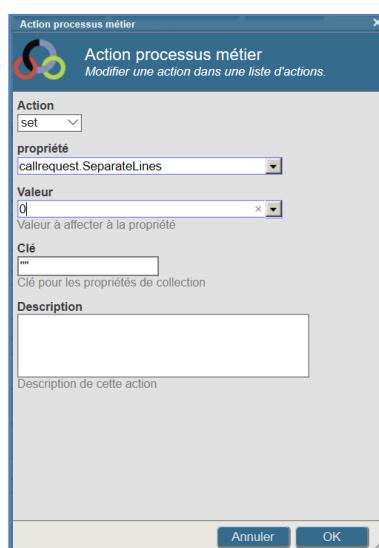
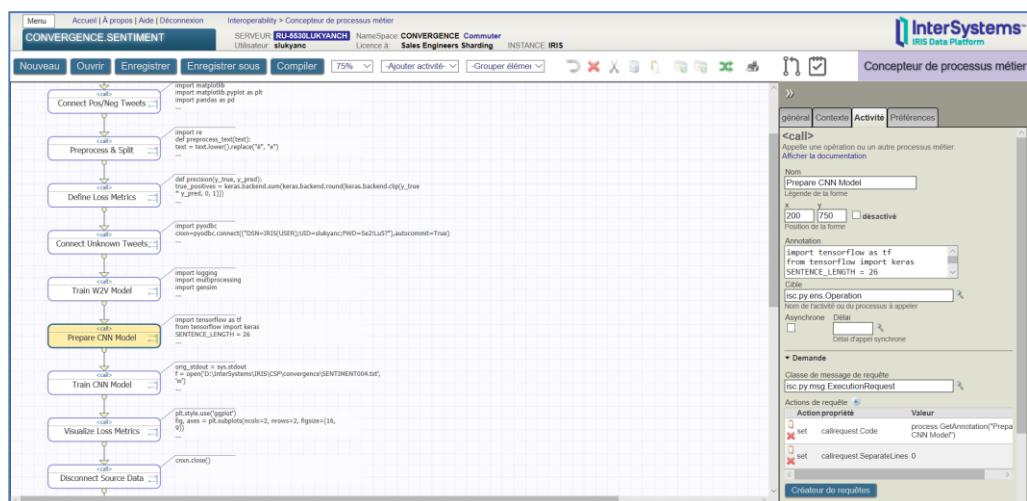
`isc.py.ens.Operation` is the ML Toolkit operation that is added to your IRIS Interoperability production. The following ML Toolkit requests are added:

1. Execute Python code via `isc.py.msg.ExecutionRequest` and get the response via `isc.py.msg.ExecutionResponse` with requested variable values as strings:

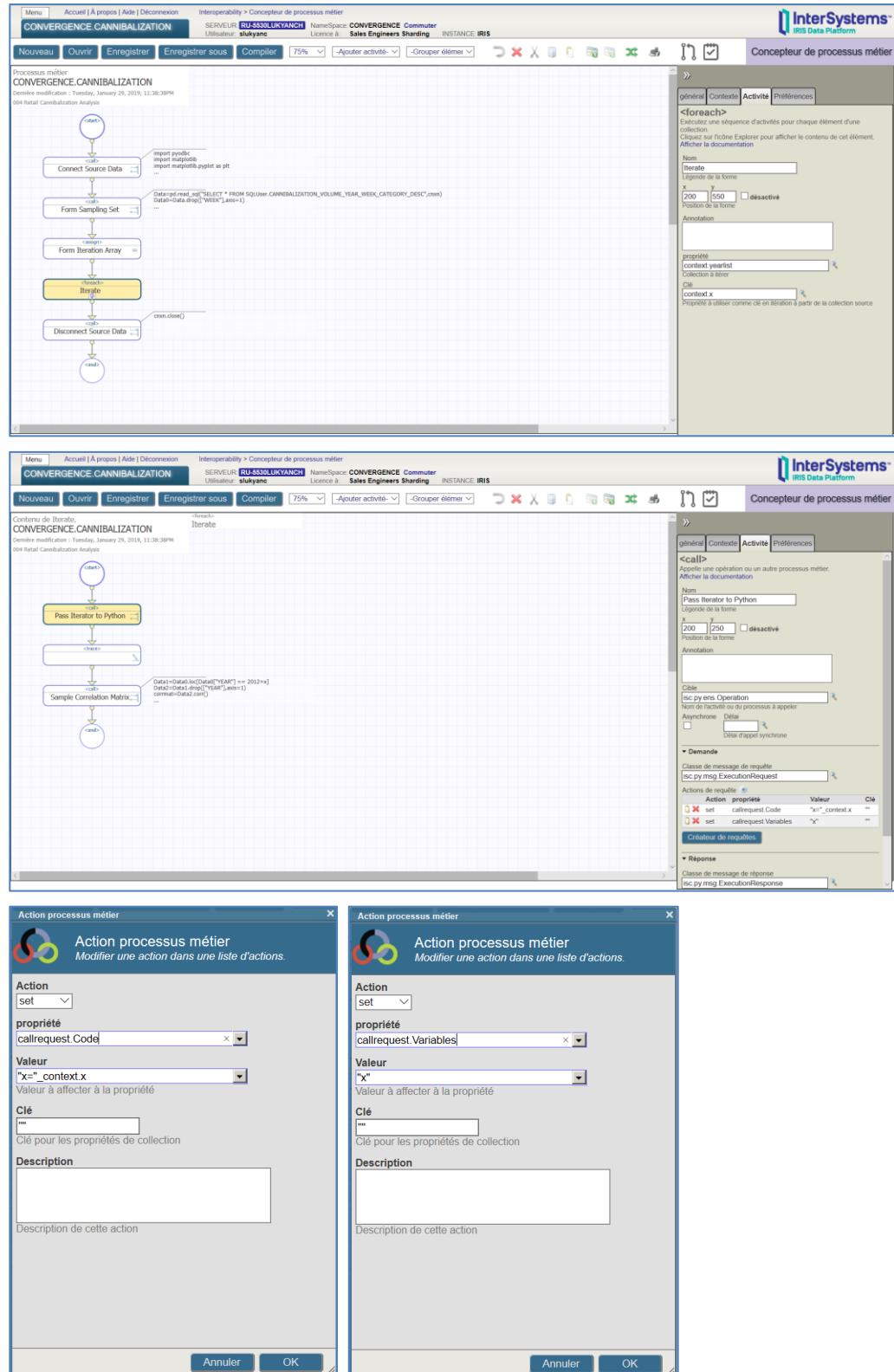




Another use of the same request (adding action to prevent line-by-line interpreting):

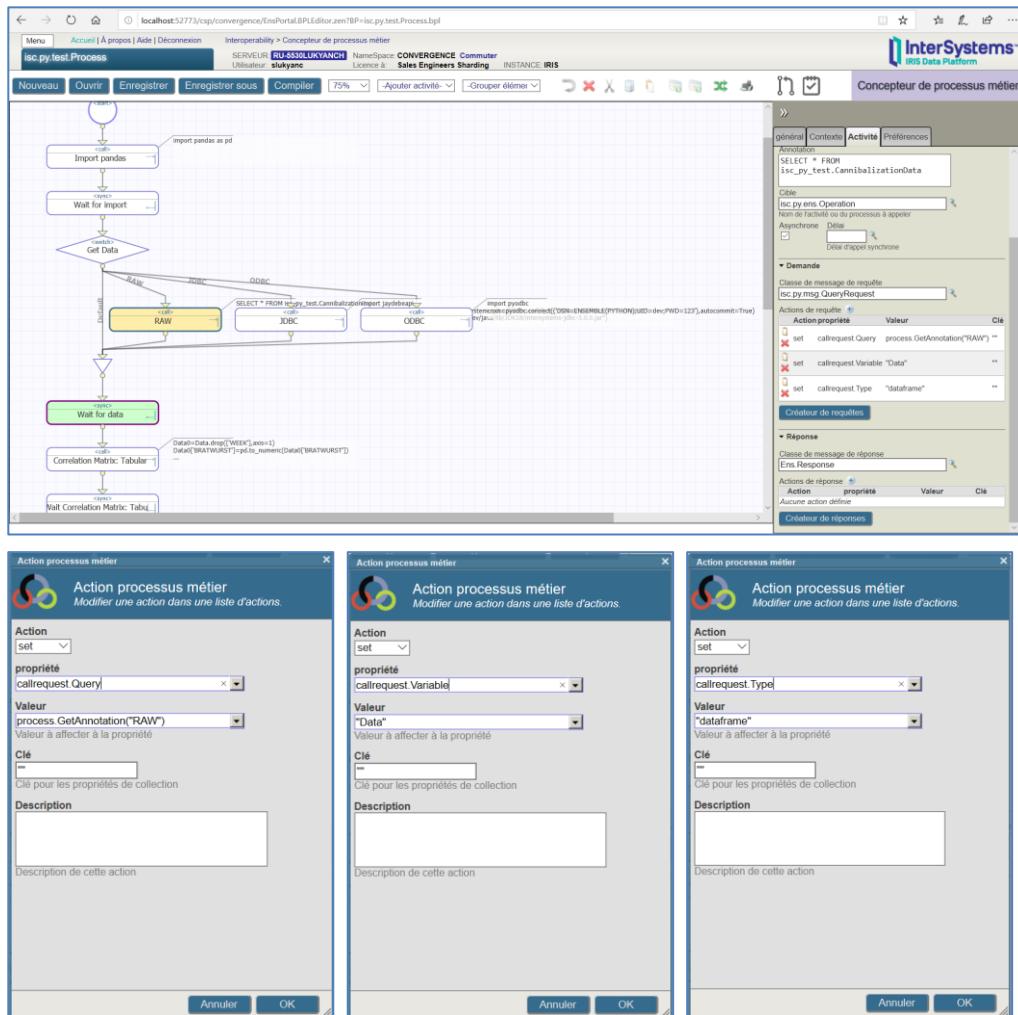


One other use of this request (adding action to pass variables between IRIS and Python contexts):

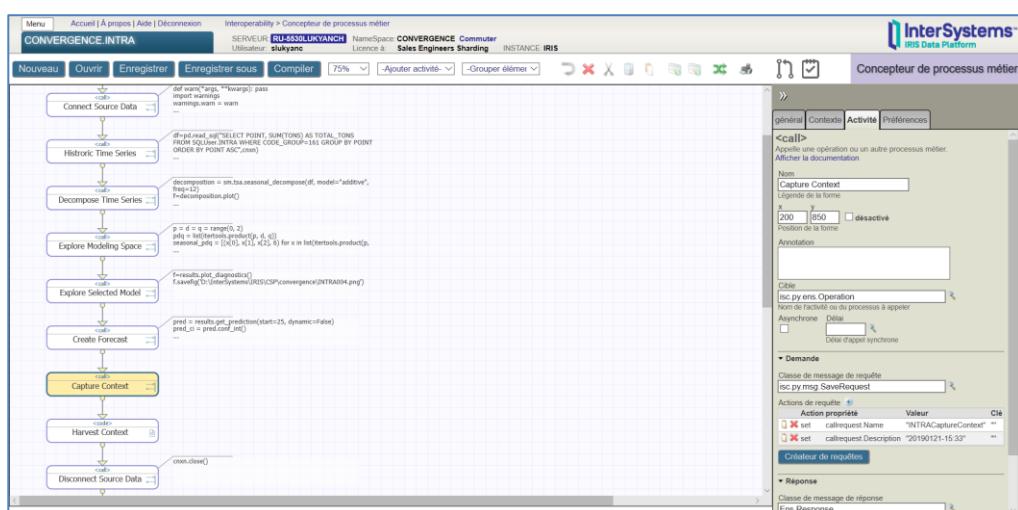


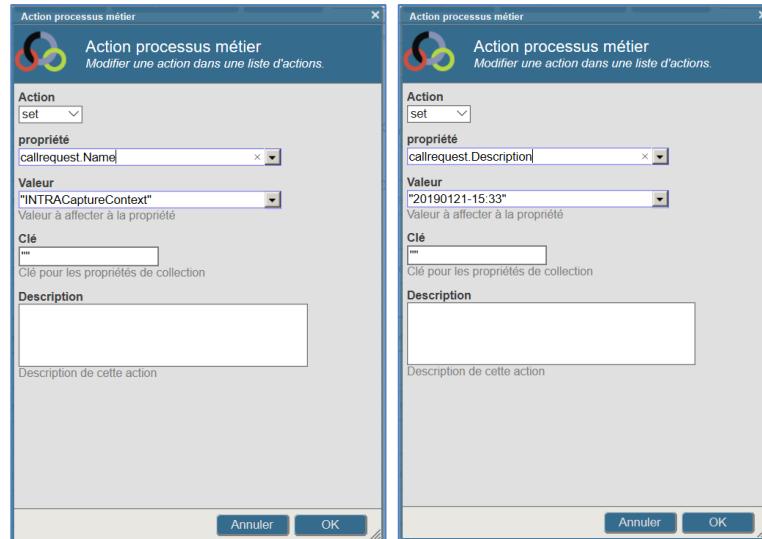
- Execute Python code via `isc.py.msg.StreamExecutionRequest` and get the response via `isc.py.msg.StreamExecutionResponse` with requested variable values as streams (everything is identical to `isc.py.msg.ExecutionRequest/isc.py.msg.ExecutionResponse` except that all code and variable values are streams)

3. Transfer data into Python from an SQL query with `isc.py.msg.QueryRequest` and get the response via `Ens.Response`



4. Save a Python context with `isc.py.msg.SaveRequest` and get the response via `Ens.StringResponse` with the context ID:





## 5. Restore a Python context with `isc.py.msg.RestoreRequest`

Action processus métier

**Action**: set

**propriété**: callrequest.ContextId

**Valeur**: "34"

Valeur à affecter à la propriété

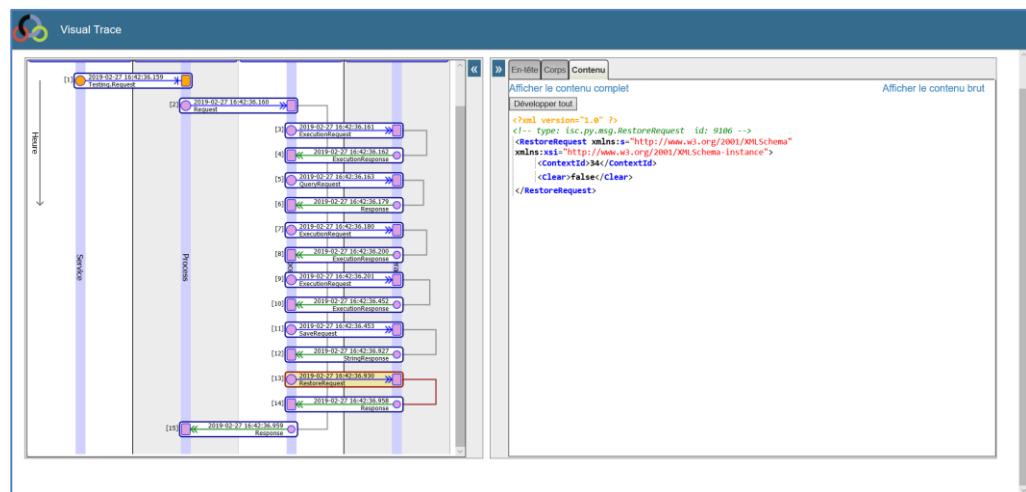
**Clé**: ""

Clé pour les propriétés de collection

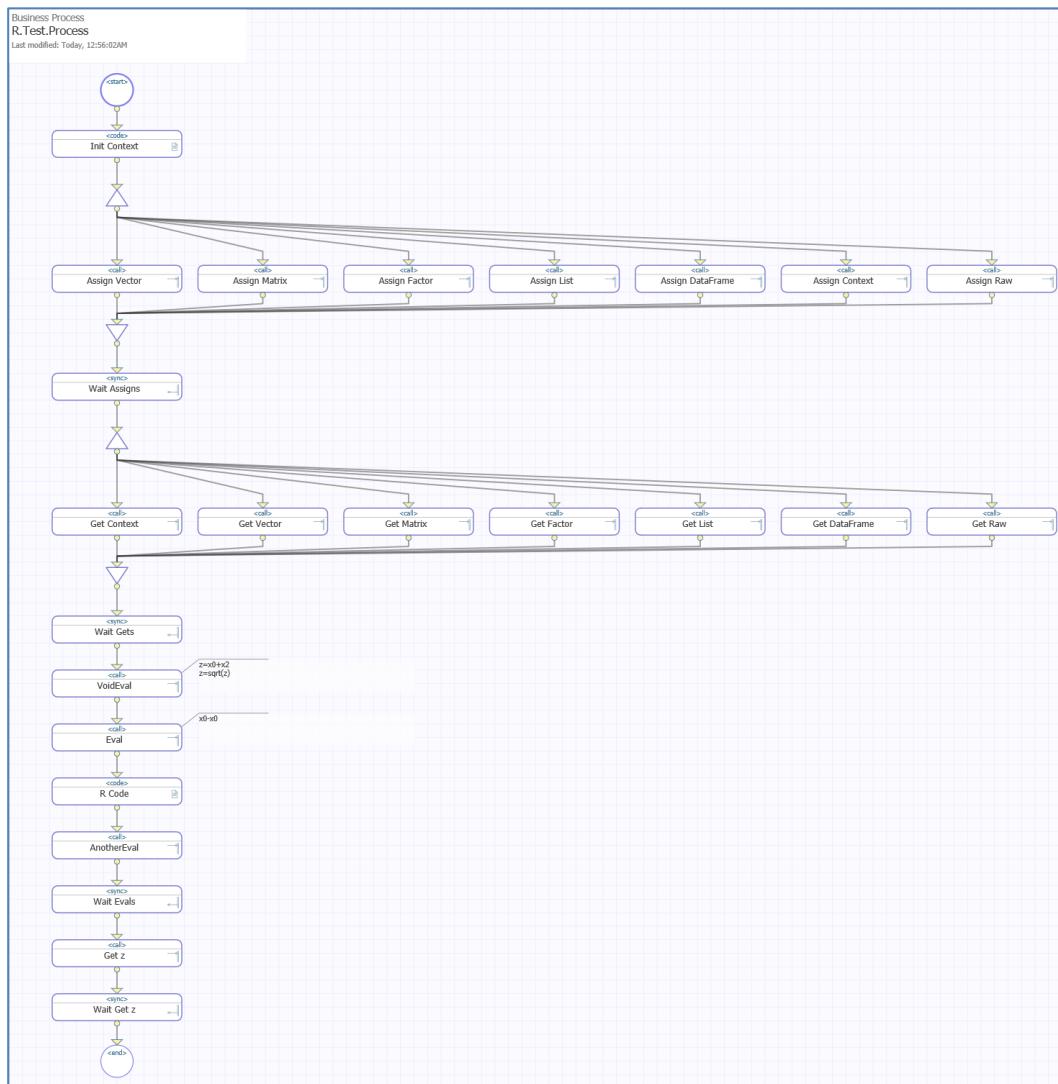
**Description**

Description de cette action

Annuler OK

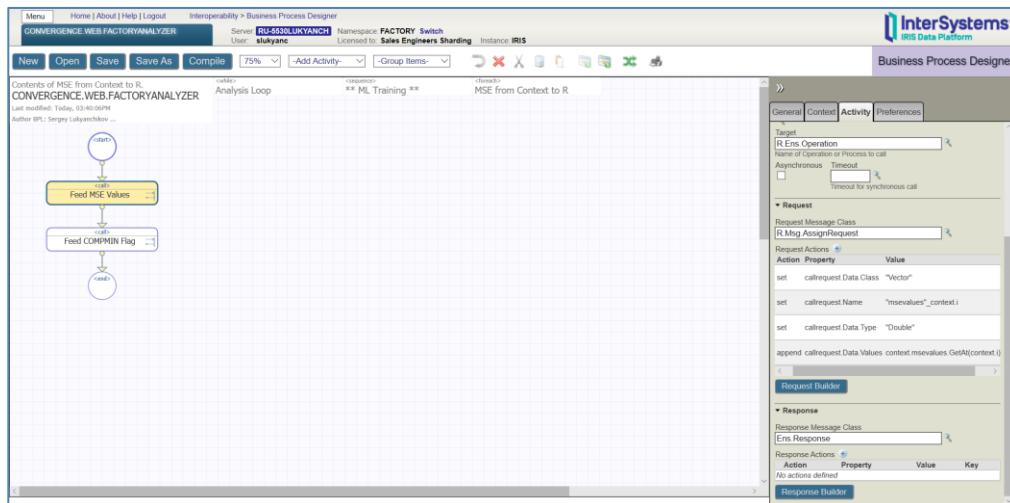


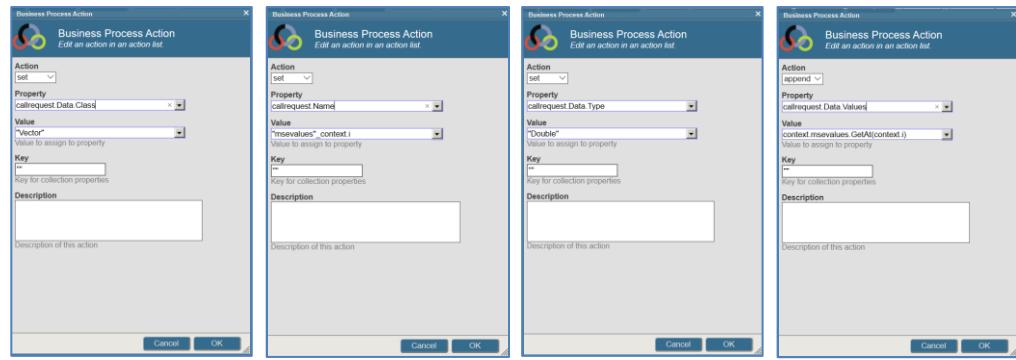
### 4.2.3. R Gateway



R.Ens.Operation is the ML Toolkit operation that is added to your IRIS Interoperability production. The following ML Toolkit requests are added:

1. Assign values to a named variable in R via R.Msg.AssignRequest:





2. Execute an R code and evaluate in IRIS-mappable terms code's named variables via `R.Msg.EvalRequest` and get the response via `R.Msg.EvalResponse`:

The screenshot shows the Business Process Designer interface with a process flow diagram. The flow starts with an 'Initial' state, followed by an 'AssignDataset' action, then a 'TrainModel' action, and finally a 'TestModel' action. Annotations in the R code sections of the actions contain IRIS-mappable terms code. A detailed view of the 'Activity' tab of the R.Msg.EvalRequest configuration dialog is shown on the right, displaying the R code and its execution context.

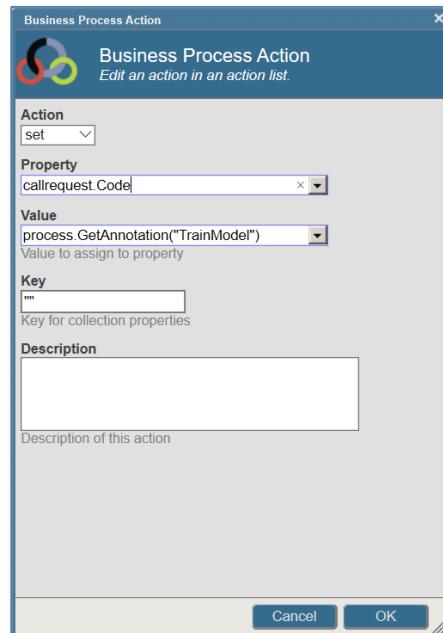
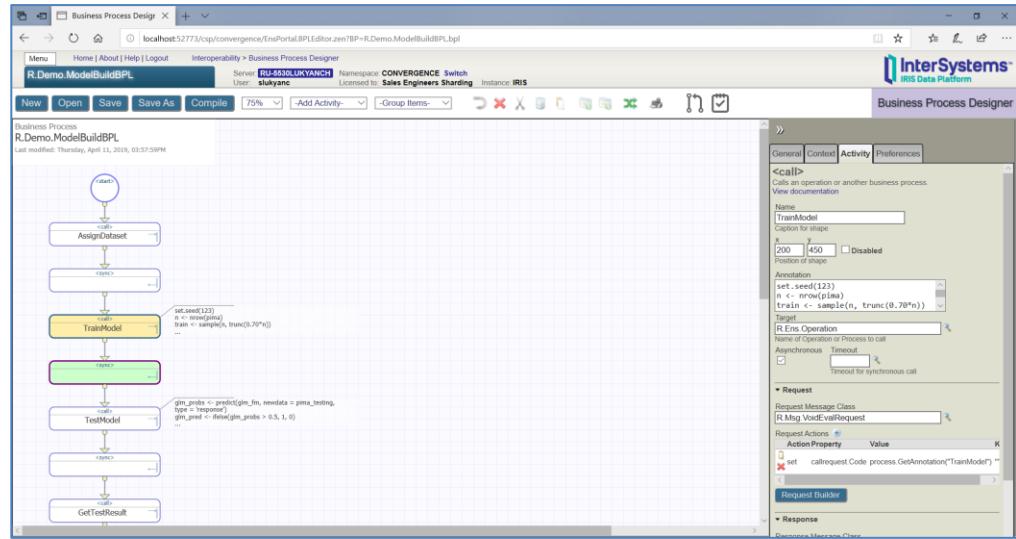
**Business Process Action Dialog (Bottom):**

```

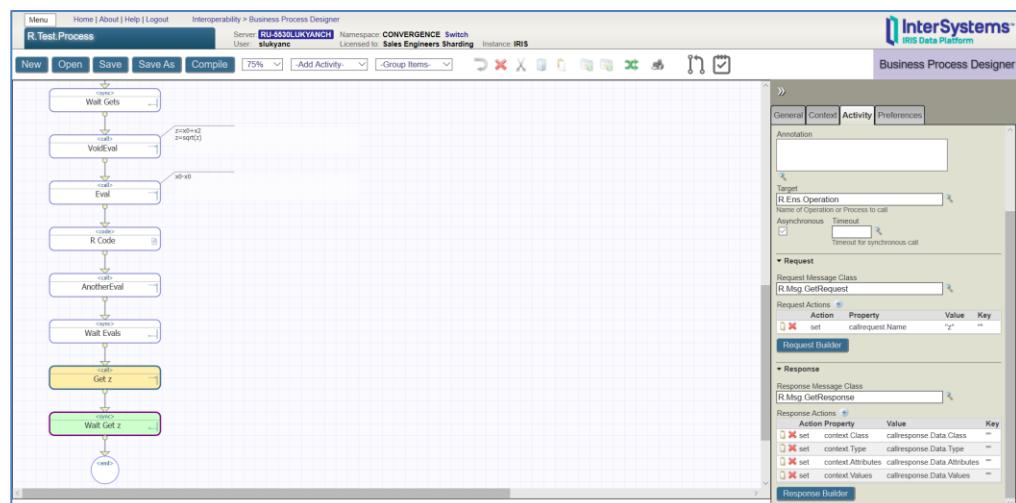
Action: set
Property: callrequest.Code
Value: process.GetAnnotation("TestModel")
Key: ...
Description: ...

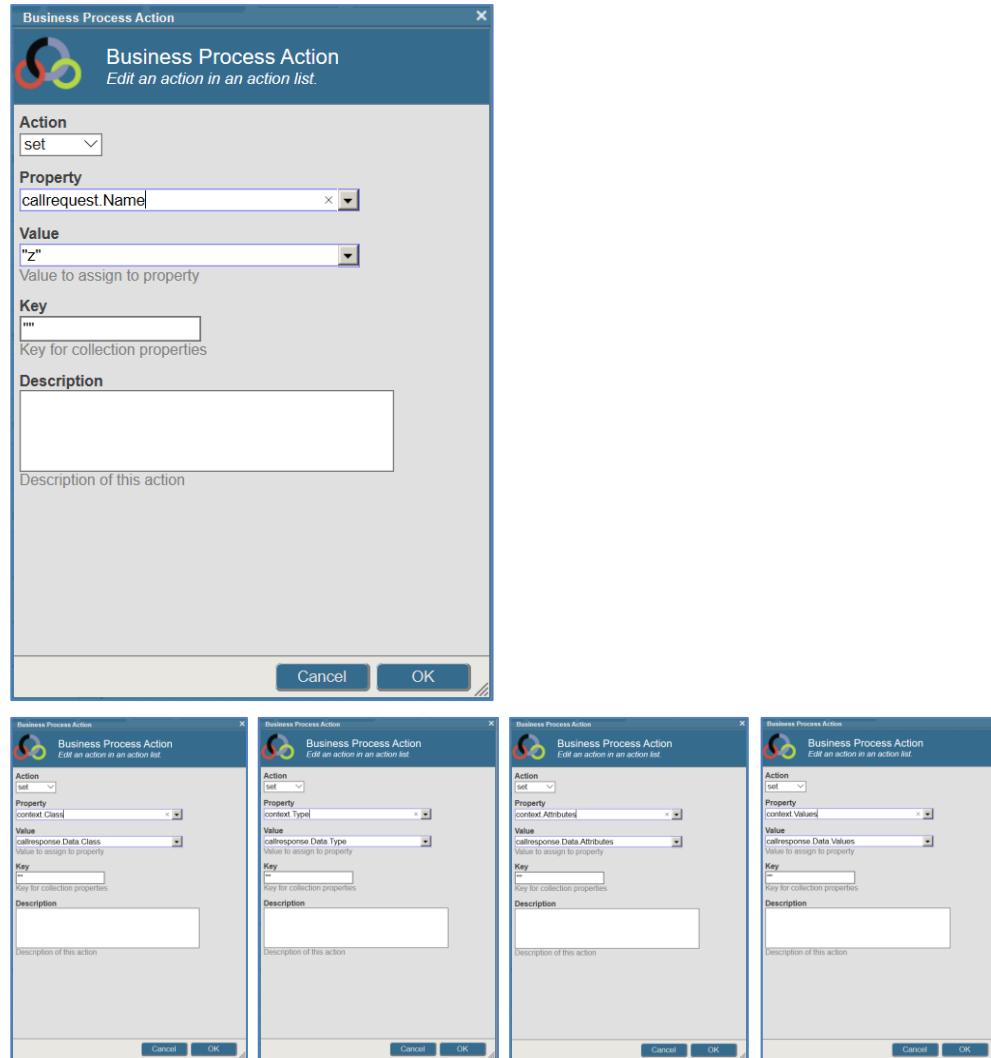
```

3. Execute an R code without evaluating in IRIS-mappable terms code's named variables with `R.Msg.VoidEvalRequest`:

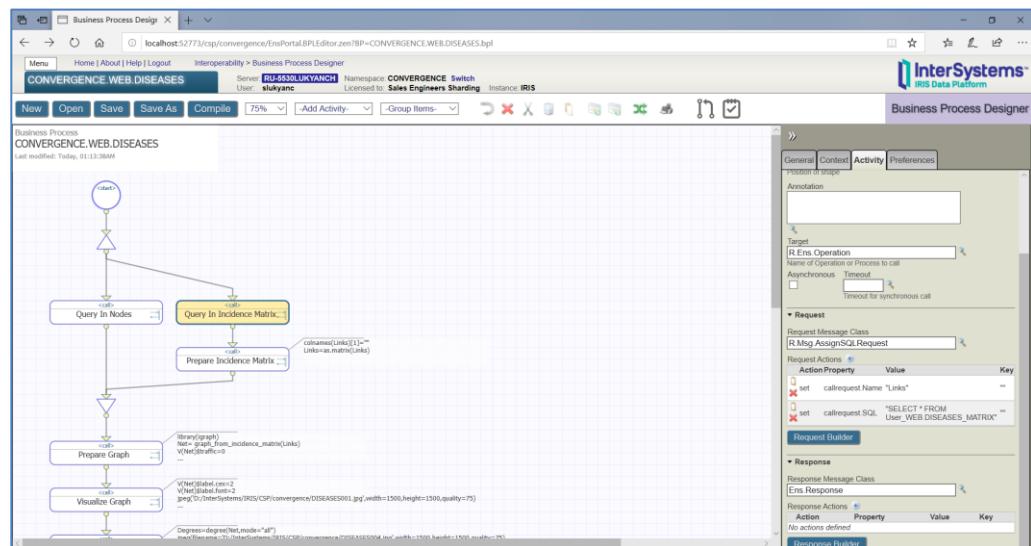


4. Retrieve values from a named variable in R with `R.Msg.GetRequest` and get the response via `R.Msg.GetResponse`:





5. Execute an SQL statement and assign its resultset as dataframe to a named variable in R via `R.Msg.AssignSQLRequest`:



**Business Process Action**  
Edit an action in an action list.

Action: set

Property: callrequest.Name

Value: "Links"

Key: ...

Description: Description of this action

**Business Process Action**  
Edit an action in an action list.

Action: set

Property: callrequest.SQL

Value: "SELECT \* FROM User\_WEB.DISEASES\_M"

Key: ...

Description: Description of this action

Cancel OK

Cancel OK

## 5. Design-Phase Tools

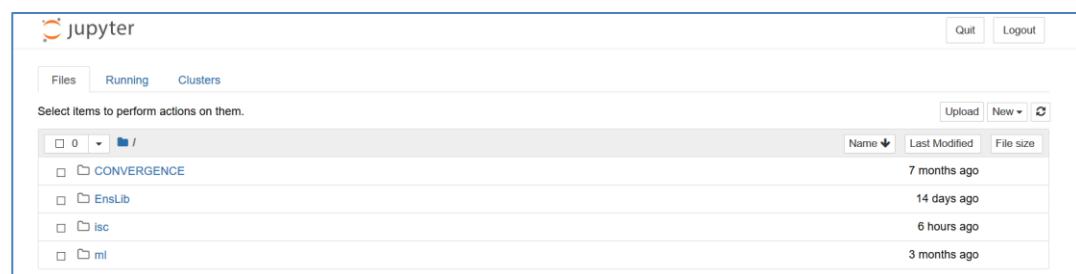
### 5.1. Jupyter Notebook

#### 5.1.1. General

[Jupyter Notebook](#) is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Our extension allows you to browse and edit InterSystems IRIS BPL processes as Jupyter notebooks. Note that currently default Python 3 executor is used. Our extension assumes that annotations contain Python code and uses activity names as preceding read-only headings.

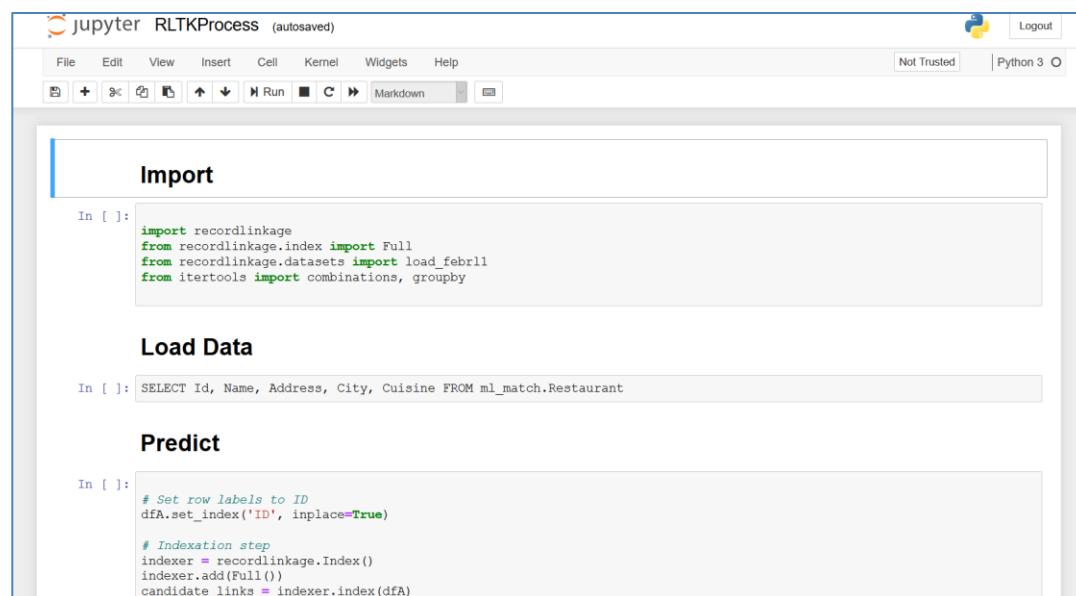
#### 5.1.2. Screenshots

##### Process Explorer



The screenshot shows the Jupyter Process Explorer interface. At the top, there are tabs for 'Files', 'Running', and 'Clusters'. Below the tabs, a message says 'Select items to perform actions on them.' A file tree is displayed under a folder named '/'. The tree includes 'CONVERGENCE', 'EnsLib', 'isc', and 'ml'. To the right of the tree, there is a table showing file details: 'Name' (with a dropdown arrow), 'Last Modified', and 'File size'. The table lists four files: '7 months ago', '14 days ago', '6 hours ago', and '3 months ago'.

##### Process Editor



The screenshot shows the Jupyter Process Editor interface. At the top, it says 'jupyter RLTKProcess (autosaved)' with a Python 3 icon and 'Logout' button. Below the title bar are menu options: File, Edit, View, Insert, Cell, Kernel, Widgets, Help. On the right, there are buttons for 'Not Trusted' and 'Python 3'. The main area contains three code cells:

- Import:** In [ ]:  

```
import recordlinkage
from recordlinkage.index import Full
from recordlinkage.datasets import load_febrl1
from itertools import combinations, groupby
```
- Load Data:** In [ ]:  

```
SELECT Id, Name, Address, City, Cuisine FROM ml.match.Restaurant
```
- Predict:** In [ ]:  

```
Set row labels to ID
dfa.set_index('ID', inplace=True)

Indexation step
indexer = recordlinkage.Index()
indexer.add(Full())
candidate_links = indexer.index(dfa)
```

#### 5.1.3. Installation

- You will need [InterSystems IRIS 2019.2 or higher](#)
- Install PythonGateway v0.8 or higher (only `isc.py.util.Jupyter` and `isc.py.ens.ProcessUtils` are required)
- Install IRISNative for Python: `pip install <IRIS>\dev\python\irisnative-* .whl`
- Install Jupyter: `pip install jupyter`

- Check [`jupyter notebook config.py`](#). It assumes the following defaults for IRIS connection:
  - `host`: localhost
  - `port`: 51773
  - `namespace`: USER
  - `user`: \_SYSTEM
  - `password`: SYS
- If you need other connection parameters values, modify `jupyter_notebook_config.py`. For example, to connect to InterSystems IRIS instance on port 51776 you will need to add this line to the bottom of `jupyter_notebook_config.py`: `c.MLContentsManager.port = 51776`
- Open OS bash in the folder with `jupyter_notebook_config.py` and `MLContentsManager.py` and start Jupyter with: `jupyter notebook`

#### 5.1.4. Notes

- Ignore “Checkpoint failed” warning on save

## 6. Convergent Analytics Showcases

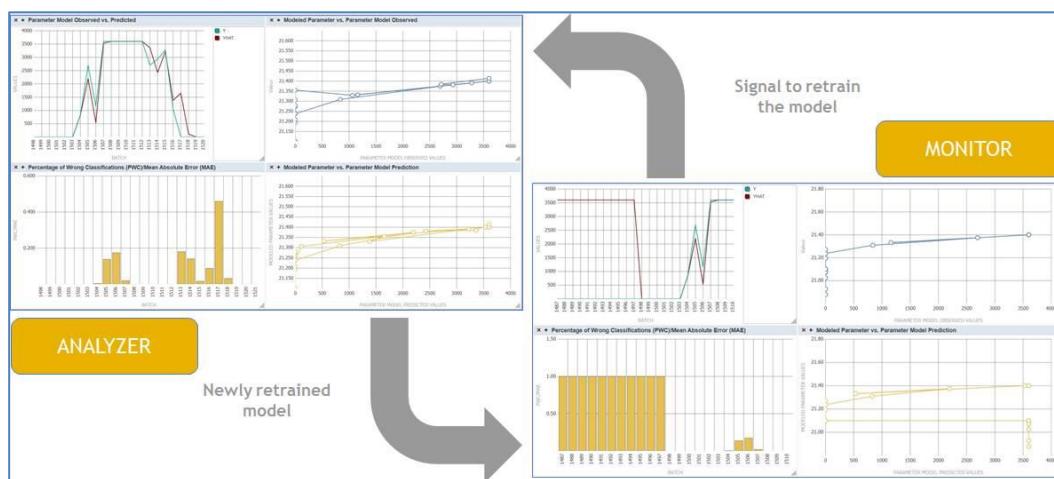
### 6.1. Featured Showcases: Robotized ML

Toolsets: Python Gateway, R Gateway

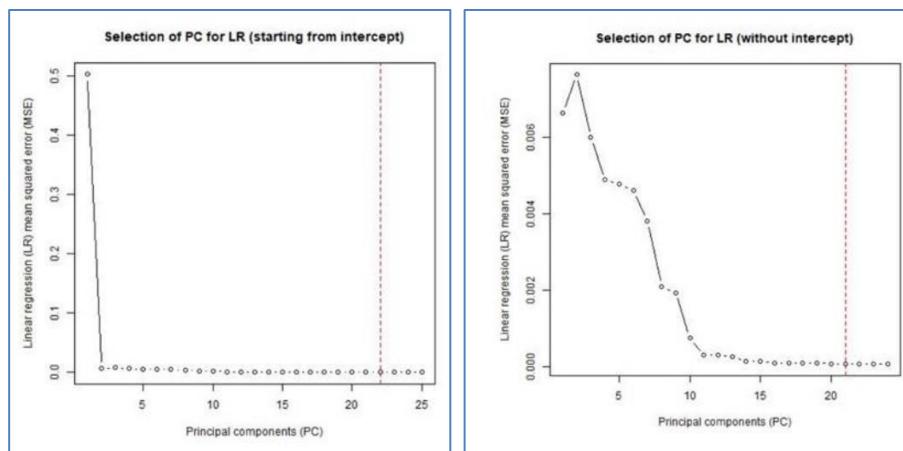
Algorithms: PCA, Linear Regression

#### 6.1.1. Background

Robotized ML: we imitate data generated by a piece of equipment by forming standard-size batches of records from a file containing a large volume of previously recorded meter readings. Those batches are being read into a buffer (if the buffer is not full), and an ML model trained on the records in the buffer is used to calculate predictions (denoted YHAT) of the target variable (denoted Y). Prediction accuracy is being monitored and if it drops below a certain threshold, the model is retrained:



The model retraining applies a light optimization technique: from a PCA projection of the original data batch, only the principal components that secure a minimum MSE are taken into the linear regression model:



#### 6.1.2. Implementation

The showcase comprises four IRIS Interoperability processes:

- GENERATOR: produces a file with the next batch of meter reading records if the previous file was imported by BUFFER

- BUFFER: imports the file produced by GENERATOR into a buffering queue if the buffer has capacity to receive another batch of records
- ANALYZER: applies an ML model to records from the buffering queue (in which case it deletes the processed records from the buffering queue to the monitoring queue), or trains/retrains the ML model (in which case it copies the records from the buffering queue and uses those records as a training sample)
- MONITOR: maintains the monitoring queue (by receiving new processed records from ANALYZER and deleting the oldest processed records from the monitoring queue), controls accuracy of the ML model and triggers its retraining if it drops below a given threshold

The showcase comprises three IRIS Analytics dashboards that refresh automatically:

- Buffer Monitor: visualizes dynamics of the data in the buffering queue
- Analyzer Monitor: visualizes the data last used for model retraining and the model validation accuracy
- Monitor Monitor: visualizes dynamics of the data and the accuracy of the predictions in the monitoring queue

The ML model used in this showcase is built to predict whether the reading of a specific sensor is above (1) or below (0) the buffer queue average – the main purpose of this is to make the model as vulnerable as possible to changes in the sensor data in the incoming record batches and to trigger frequent model retraining.

### 6.1.3. Walkthrough

- **ACTION:** before starting the showcase, if you would like to limit the disk space consumption on your computer, execute the following commands in IRIS Terminal:
  - zn "%SYS"
  - do ^JOURNAL

At the journaling management utility menu startup, choose the following options:

- “2 – Stop Journaling”
- Answer “Y” if asked whether you would like to stop journaling now

If the showcase has been already started, left working and then stopped – all without a prior journaling stopped, in addition to the above options please add:

- “6 – Purge Journal Files”
- Choose sub-option “1 – Purge any journal NOT required for transaction rollback or crash recovery”

**NOTE:** if the computer you are using for the showcase is not under your administration, you are requested to consult the administrator before proceeding with the above actions.

- **ACTION:** before starting the showcase, copy the file `provide.csv` to a folder chosen as the showcase working folder (**NOTE:** the folder you choose as working must be accessible for IRIS, therefore it is optimal to select the working folder under your IRIS installation path).
- For a more efficient control of the data queues and the model accuracy, use the following SQL queries (we recommend opening a separate IRIS Management Portal session for each of the below queries to simplify usage):
  - SELECT  
`BATCH, COUNT(*) AS RECORDS, ABS(ROUND(MAX(STAMP))-MIN(STAMP),2) AS THRUPUT`

```

 FROM User_WEB.FACTORYBUFFER
 GROUP BY BATCH
 ORDER BY BATCH ASC
 • SELECT
 BATCH, COUNT(*) AS RECORDS
 FROM User_WEB.FACTORYANALYZER
 GROUP BY BATCH
 ORDER BY BATCH ASC
 • SELECT
 BATCH, COUNT(*) AS RECORDS
 FROM User_WEB.FACTORYMONITOR
 GROUP BY BATCH
 ORDER BY BATCH ASC
 • SELECT DISTINCT Y, YHAT, COUNT(*) FREQUENCY
 FROM User_WEB.FACTORYMONITOR
 GROUP BY Y, YHAT
 • SELECT SUM(ABS(Y-YHAT))/COUNT(*) AS PWC FROM
 User_WEB.FACTORYMONITOR

```

- To allow starting the showcase, create in the working folder the `go.txt` empty text file. To stop the showcase (all its four BPL processes) at any moment, delete the `go.txt` file from the working folder.
- **NOTE:** the order, in which you start the BPL processes is not important (they will all wait for enough action to be taken by each of them to start functioning). However, we do recommend the order below during demos to make your story easier to understand for your audience.
- Start GENERATOR process (all processes in this showcase are started by test-sending them a generic `Ens.Request` message). During demos, do not wait till a visual trace is generated – all processes in this showcase are supposed to be working indefinitely long.
- Start BUFFER process and wait till the buffering queue is filled (use the first from the SQL queries above).
- Start ANALYZER process and wait till its data table get filled (use the second from the SQL queries above), then wait till the first training of the ML model is finished (for that, use queue monitoring in IRIS Interoperability to capture the moment that `isc.py.ens.Operation` finishes its work and check the graphical files generated by the ML model in the working folder).
- Start MONITOR process and check that the monitoring queue gets filled with at least 12 record batches (use the third from the SQL queries above).
- Use the fourth and the fifth from the SQL queries above to control the confusion matrix and the percentage of wrong classifications (PWC) of the ML model.
- Start Buffer Monitor, Analyzer Monitor and Monitor Monitor dashboards.

## 6.2. Featured Showcases: Robotized Sentiment Analysis (English)

Toolsets: Python Gateway

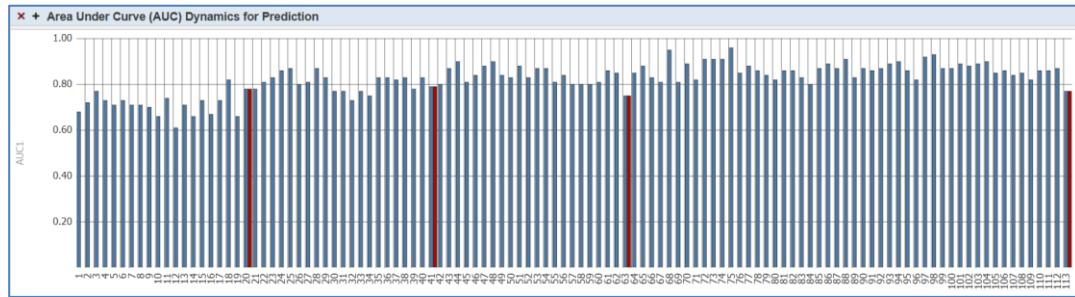
Algorithms: Word2Vec, RNN

Idea: <https://stackabuse.com/python-for-nlp-movie-sentiment-analysis-using-deep-learning-in-keras>

Data: <https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>

### 6.2.1. Background

Robotized Sentiment Analysis: we imitate data produced by a social network subscription by forming standard-size batches of records from files containing labeled (positive and negative) and unlabeled posts in Twitter (tweets). Those batches are being read into respective buffers (if the buffers are not full), and an ML model trained on the records in the labeled buffers is used to calculate the probability of an unlabeled tweet to be positive. Prediction accuracy is being evaluated on the labeled records used for training; if it does not reach a certain accuracy threshold, the model is retrained:



The resulting sentiment classification is visualized using a dashboard:

| Text (Excerpt)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | YHAT                                                                                 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| 'P' (or Club-P) should really be called 'L' for lame. Every festival has a disappointment and this is the one that fails to live up to its much-hyped (SPOILERS included) This film surely is the best Amicus production I've seen so far (even though I still have quite a few to check out). The House th... good script or director couldn't fix.<br /><br />The original 'Poseidon Adventure' was a story of human courage triumphing against terrific odds.<br /><br />The kind of movie that demands too much from the audience, yet to disappoint most of them. I think of the Mexican people who (either in a<br /><br />Very dull, laborious adaptation of Amis's amusing satire. The hero is portrayed not as a likable loser but a merely callous cretin. Most A college professor is working on creating zombies and, wouldn't you know it, things go horribly wrong. Next thing you know, the whole campus is over A delightful piece of cinema storytelling in a simple but effective way. Cinema after all is a visual media and Igor used its full potential. A young A few years ago I saw The Scent of Green Papaya by the same director. My feelings about both films are in fact the same: beautifully shot, but tern A friend warned me that this was the worst movie he had ever seen. I was curious, because it had developed a bit of a cult status. I like a lot of od A funny thing happened to me while watching Mosquito: on the one hand, the hero is a deaf-mute and the director is totally unable to make us underst A lot of death happens in the wild. You don't need to be a rocket scientist to figure that out! But does it need to be the focus of a nature document | 0.59<br>0.68<br>0.75<br>0.71<br>0.66<br>0.61<br>0.94<br>0.61<br>0.78<br>0.60<br>0.86 |

| Text (Excerpt)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | YHAT                                                                                 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| 'The Omen 4: The Awakening' is a made-for-television sequel to the original 'The Omen' film. Instead of Satan possessing the body of a little boy, he... may contain spoilers.<br /><br />Clearly, who ever made this film must have had a lot of connections. I just can't see it any other way. What really...<br /><br />I am a big-time horror/sci-fi fan regardless of budget, but after watching countless horror movies late night on cable and video, this ha...<br /><br />Very slow, plodding movie with a confusing story line. The movie's only hope of keeping the audience interested is the gratuitous nudity A group of friends receive word from a pal who has found gold in an old mine shaft nearby an ancient wasteland town of Sutterville. Despite A meltdown at a nuclear power plant causes a majority of people to turn lethal, rot-faced, shambling zombies who naturally go on a grisly rampage A remake of the 1916 silent film, based on the 1909 novel by Maurice Leblanc. The detective series would be made into numerous plays, films and TV se A young woman comes to the home town of his husband after he passed away in an accident. She barely settles down in this small town, but shortly afte Absolutely one of the worst movies I've seen in a long time! It starts off badly and just deteriorates. Katherine Heigl is woefully miscast in a Loli After seeing the trailer it was an easy decision not to see this film. I mean, I don't care for stupid stoner comedies. I'm sure it was also an easy ALERT. This review contains major SPOILERS. Do not read on if you plan to see this film.<br /><br />Judging from the amount of votes this got (my vot | 0.28<br>0.34<br>0.21<br>0.42<br>0.35<br>0.31<br>0.12<br>0.27<br>0.07<br>0.23<br>0.25 |

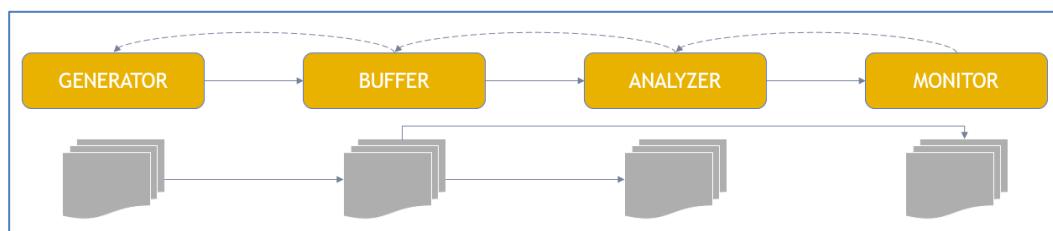
### 6.2.2. Implementation

The showcase comprises four IRIS Interoperability processes:

- GENERATOR: as previously generated files get consumed by the other processes, generates new files with input data (labeled – positive and negative tweets – as well as unlabeled tweets)

- **BUFFER:** as already buffered records are consumed by the other processes, reads new records from the files created by GENERATOR and deletes the files after having read records from them
- **ANALYZER:** consumes records from the unlabeled buffer and applies to them the trained RNN (recurrent neural network), transfers the “applied” records with respective “probability to be positive” values added to them, to the monitoring buffer; consumes records from labeled (positive and negative) buffers and trains the neural network based on them
- **MONITOR:** consumes records processed and transferred to its buffer by ANALYZER, evaluates the classification error metrics demonstrated by the neural network after the last training, and triggers new training by ANALYZER

The agent-based system of processes used in this showcase can be illustrated as follows:



The ML mechanisms used in this showcase are a file with “frozen” text vectorization results and a recurrent neural network to model sentiment; the training dataset is grown slowly to trigger multiple model training runs.

### 6.2.3. Walkthrough

- **ACTION:** before starting the showcase, if you would like to limit the disk space consumption on your computer, execute the following commands in IRIS Terminal:
  - `zn "%SYS"`
  - `do ^JOURNAL`

At the journaling management utility menu startup, choose the following options:

- “2 – Stop Journaling”
- Answer “Y” if asked whether you would like to stop journaling now

If the showcase has been already started, left working and then stopped – all without a prior journaling stopped, in addition to the above options please add:

- “6 – Purge Journal Files”
- Choose sub-option “1 – Purge any journal NOT required for transaction rollback or crash recovery”

**NOTE:** if the computer you are using for the showcase is not under your administration, you are requested to consult the administrator before proceeding with the above actions.

- **ACTION:** before starting the showcase, copy the files `provide*.csv` and `glove.6B.100d.txt` to a folder chosen as the showcase working folder (**NOTE:** the folder you choose as working must be accessible for IRIS, therefore it is optimal to select the working folder under your IRIS installation path).
- For a more efficient control of the data queues and the model accuracy, use the following SQL queries (we recommend opening a separate IRIS Management Portal session for each of the below queries to simplify usage):
  - `SELECT`

- ```

BATCH, COUNT(*) AS RECORDS, ABS(ROUND(MAX(STAMP)-
MIN(STAMP),2)) AS THRUPUT
FROM User_WEB.SENTIMENTENBUFFERUNK (or else, *POS and *NEG)
GROUP BY BATCH
ORDER BY BATCH ASC
• SELECT
    BATCH, COUNT(*) AS RECORDS
    FROM User_WEB.SENTIMENTENANALYZERPOS (also, *NEG)
    GROUP BY BATCH
    ORDER BY BATCH ASC
• SELECT
    BATCH, COUNT(*) AS RECORDS
    FROM User_WEB.SENTIMENTENMONITOR
    GROUP BY BATCH
    ORDER BY BATCH ASC
• SELECT *
    FROM User_WEB.SENTIMENTENANALYZERAUC

```
- To allow starting the showcase, create in the working folder the `go.txt` empty text file. To stop the showcase (all its four BPL processes) at any moment, delete the `go.txt` file from the working folder.
 - **NOTE:** the order, in which you start the BPL processes is not important (they will all wait for enough action to be taken by each of them to start functioning). However, we do recommend the order below during demos to make your story easier to understand for your audience.
 - Start GENERATOR process (all processes in this showcase are started by test-sending them a generic `Ens.Request` message). During demos, do not wait till a visual trace is generated – all processes in this showcase are supposed to be working indefinitely long.
 - Start BUFFER process and wait till the buffering queues are filled (use the first from the SQL query templates above).
 - Start ANALYZER process and wait till its data tables get filled (use the second and the fourth from the SQL query templates above), then wait till the first training of the ML models is finished (for that, use queue monitoring in IRIS Interoperability to capture the moment that `isc.py.ens.Operation` finishes its work and check the text/graphical files generated by the ML models in the working folder).
 - Start MONITOR process and check that its data tables get filled (use the third and the fourth from the SQL query templates above).

Start Buffer Monitor, Analyzer Monitor and Monitor Monitor dashboards.

6.3. Featured Showcases: Robotized Sentiment Analysis (Russian)

Toolsets: Python Gateway

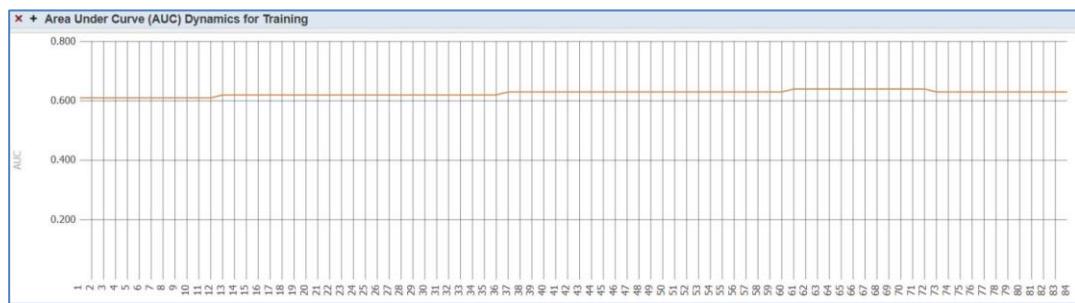
Algorithms: Word2Vec, CNN

Idea: <https://github.com/sismetanin/sentiment-analysis-of-tweets-in-russian>

Data: <http://study.mokoron.com/>

6.3.1. Background

Robotized Sentiment Analysis: we imitate data produced by a social network subscription by forming standard-size batches of records from files containing labeled (positive and negative) and unlabeled posts in Twitter (tweets). Those batches are being read into respective buffers (if the buffers are not full), and an ML model trained on the records in the labeled buffers is used to calculate the probability of an unlabeled tweet to be positive. Prediction accuracy is being evaluated on the labeled records used for training; if it does not reach a certain accuracy threshold, the model is retrained:



The resulting sentiment classification is visualized using a dashboard:

User	Tweet Content	YHAT
#Kredit Банки выдающие кредиты без справок http://it.co/ZNIRWwkPEx	@nuzulka скажите пцп Родион Газманов - Гравитация http://it.co/DNEC6abPA	0.65
ищите История из жизни, верных и преданных пись Марат http://it.co/bfBQbLwng3 #жизнь		0.64
8 декабря в 16:00 В школе активного горожанина продолжат разбираться с платежами за жилищно-коммунальные услуги http://it.co/JTFmDHoLZ		0.66
@bab14 http://it.co/3OcV2dJU7u, бахтреисы в продучене... Какое ух тут негроцное.		0.68
@JuliaTrolol4 АХАХА ПРИ ЧЕМ ТУТ БАЛЕРИНА ВАПЩИ? ПХАХА УПРТ		0.65
@coozy_ ехахах http://it.co/PMOkvUBcI		0.55
@eniripovich мой личный друг Морган Фримене http://it.co/MFb7BaZen		0.85
Bluetooth игры на андроид http://it.co/MbSboNgor		0.63
Glamour Fashion Night 2013 (клуб 6 Насы) 12 декабря 2013 (четверг) Клубы Орла Орен Афиша http://it.co/36IJ5Almy		0.67
http://it.co/0IVbcKF1AR седло и уздача скачать		0.67

User	Tweet Content	YHAT
~ что делать с собой?		0.45
@Annya_Nik @YanikaMurygina Так и думал:[0.42
@Mashikh_tyts да так		0.43
@Norenko_ ясно) что делать?		0.42
@redutusowl @Zirchik хотя не, норм		0.44
@WEBelarus ну и я что не сильно вот		0.43
@_r_a_l_no... как же... и без тебя...		0.43
А что такое... долг до этого что-то плохое?		0.43
Голова болит(никогда больше не буду пить так)		0.43
Жена говорит, что у меня гипертонии, но я ведь даже не женат.		0.41
какая же... ениская погода?"^		0.44
НЕ ИМЕЕТ Смысла...		0.43
Не спасить от исключения тому, кто боится.		0.41

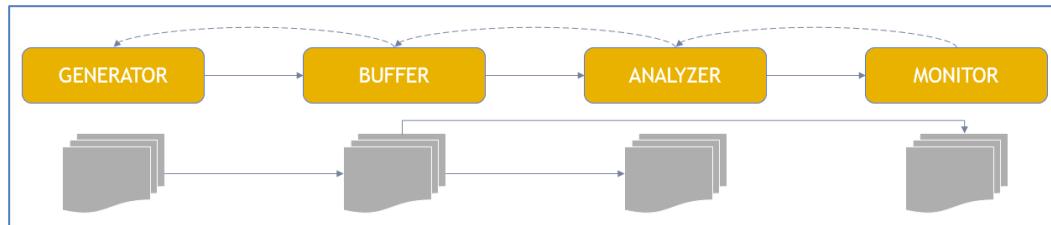
6.3.2. Implementation

The showcase comprises four IRIS Interoperability processes:

- GENERATOR: as previously generated files get consumed by the other processes, generates new files with input data (labeled – positive and negative tweets – as well as unlabeled tweets)
- BUFFER: as already buffered records are consumed by the other processes, reads new records from the files created by GENERATOR and deletes the files after having read records from them

- ANALYZER: consumes records from the unlabeled buffer and applies to them the trained CNN (convolutional neural network), transfers the “applied” records with respective “probability to be positive” values added to them, to the monitoring buffer; consumes records from labeled (positive and negative) buffers and trains the neural network based on them
- MONITOR: consumes records processed and transferred to its buffer by ANALYZER, evaluates the classification error metrics demonstrated by the neural network after the last training, and triggers new training by ANALYZER

The agent-based system of processes used in this showcase can be illustrated as follows:



The ML mechanisms used in this showcase are a text vectorizer and a convolutional neural network to model sentiment; the training dataset is grown slowly to trigger multiple model training runs.

6.3.3. Walkthrough

- **ACTION:** before starting the showcase, if you would like to limit the disk space consumption on your computer, execute the following commands in IRIS Terminal:

- zn "%SYS"
- do ^JOURNAL

At the journaling management utility menu startup, choose the following options:

- “2 – Stop Journaling”
- Answer “Y” if asked whether you would like to stop journaling now

If the showcase has been already started, left working and then stopped – all without a prior journaling stopped, in addition to the above options please add:

- “6 – Purge Journal Files”
- Choose sub-option “1 – Purge any journal NOT required for transaction rollback or crash recovery”

NOTE: if the computer you are using for the showcase is not under your administration, you are requested to consult the administrator before proceeding with the above actions.

- **ACTION:** before starting the showcase, copy the files provide*.csv to a folder chosen as the showcase working folder (**NOTE:** the folder you choose as working must be accessible for IRIS, therefore it is optimal to select the working folder under your IRIS installation path).
- For a more efficient control of the data queues and the model accuracy, use the following SQL queries (we recommend opening a separate IRIS Management Portal session for each of the below queries to simplify usage):
 - SELECT
 BATCH, COUNT(*) AS RECORDS, ABS(ROUND(MAX(STAMP))-MIN(STAMP),2) AS THRUPUT
 FROM User_WEB.SENTIMENTRUBUFFERUNK (or else, *POS and *NEG)

```
GROUP BY BATCH
ORDER BY BATCH ASC
• SELECT
BATCH, COUNT(*) AS RECORDS
FROM User_WEB.SENTIMENTRUANALYZERPOS (also, *NEG)
GROUP BY BATCH
ORDER BY BATCH ASC
• SELECT
BATCH, COUNT(*) AS RECORDS
FROM User_WEB.SENTIMENTRUMONITOR
GROUP BY BATCH
ORDER BY BATCH ASC
```

- To allow starting the showcase, create in the working folder the `go.txt` empty text file. To stop the showcase (all its four BPL processes) at any moment, delete the `go.txt` file from the working folder.
- **NOTE:** the order, in which you start the BPL processes is not important (they will all wait for enough action to be taken by each of them to start functioning). However, we do recommend the order below during demos to make your story easier to understand for your audience.
- Start GENERATOR process (all processes in this showcase are started by test-sending them a generic `Ens.Request` message). During demos, do not wait till a visual trace is generated – all processes in this showcase are supposed to be working indefinitely long.
- Start BUFFER process and wait till the buffering queues are filled (use the first from the SQL query templates above).
- Start ANALYZER process and wait till its data tables get filled (use the second from the SQL query templates above), then wait till the first training of the ML models is finished (for that, use queue monitoring in IRIS Interoperability to capture the moment that `isc.py.ens.Operation` finishes its work and check the text/graphical files generated by the ML models in the working folder).
- Start MONITOR process and check that the monitoring queue gets filled with several record batches (use the third from the SQL query templates above).
- Start Buffer Monitor, Analyzer Monitor and Monitor Monitor dashboards.

6.4. 001 Sentiment Analysis

Toolsets: Python Gateway

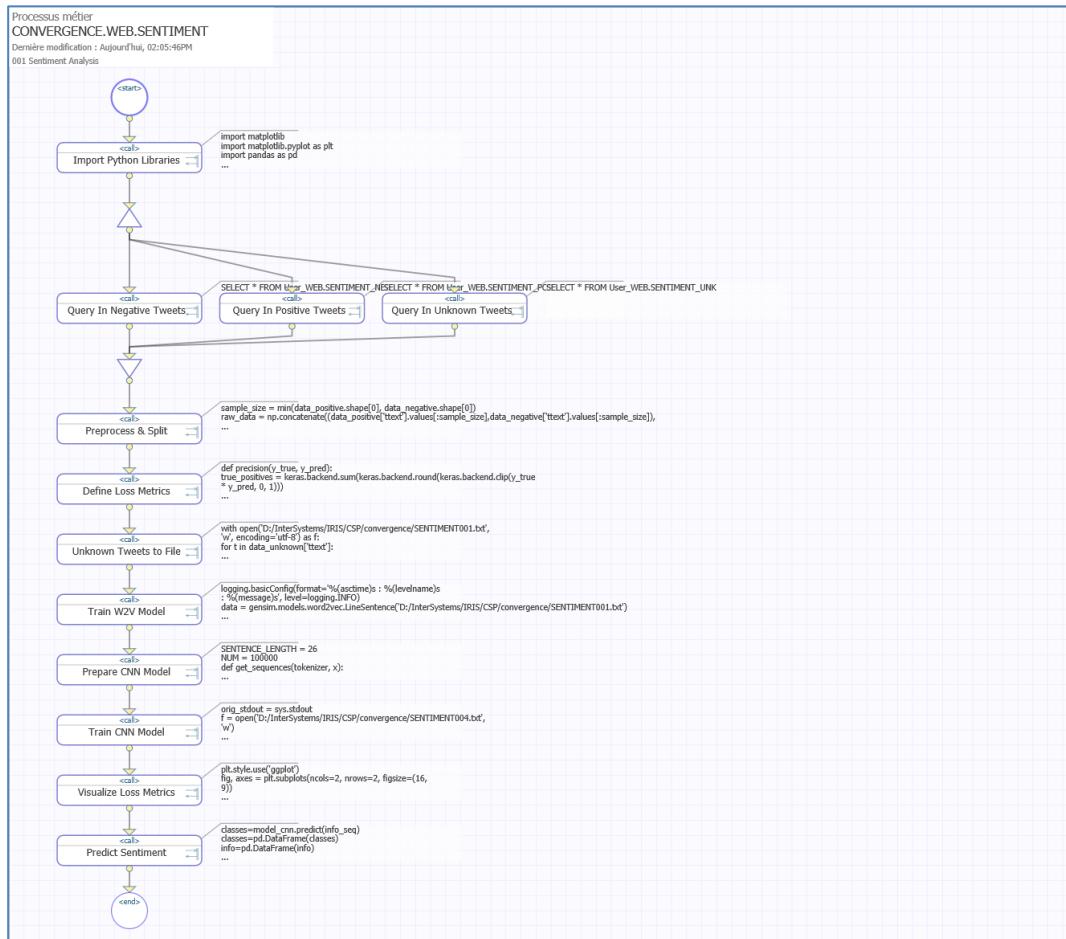
Algorithms: Word2Vec, CNN

6.4.1. Background

Sentiment analysis: we transform words in the tweets into vectors, then train a model that relies on vector representation to establish proximity of a given tweet to positive or negative tweets (based on vectorized samples of both).



6.4.2. Implementation



6.4.3. Walkthrough

- Import Python Libraries: loads the required libraries to Python context.
- Query In Negative Tweets: loads a sample of negative tweets. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
- Query In Positive Tweets: loads a sample of positive tweets.
- Query In Unknown Tweets: loads a sample of the tweets to be classified as either negative or positive.
- Preprocess & Split: adjusts texts in the negative and positive samples to ensure efficient processing via the vectorization and neural network components. **ACTION:** adjust the file paths to fit your environment.
- Define Loss Metrics: defines functions to calculate loss metrics that measure classification accuracy.
- Unknown Tweets to File: writes preprocessed texts from the unknown sample into a text file to be fed in the vectorization component.
- Train W2V Model: converts texts to numerical vectors to enable computations using the neural network.
- Prepare CNN Model: defines and initializes a convolutional neural network (CNN).
- Train CNN Model: trains and validates the CNN model.
- Visualize Loss Metrics: visualizes the progress that the loss metrics are making as epochs are being processed.

- Predict Sentiment: scores unknown tweets for their sentiment probability to be positive (scores are closer to 1) or negative (scores are closer to 0)

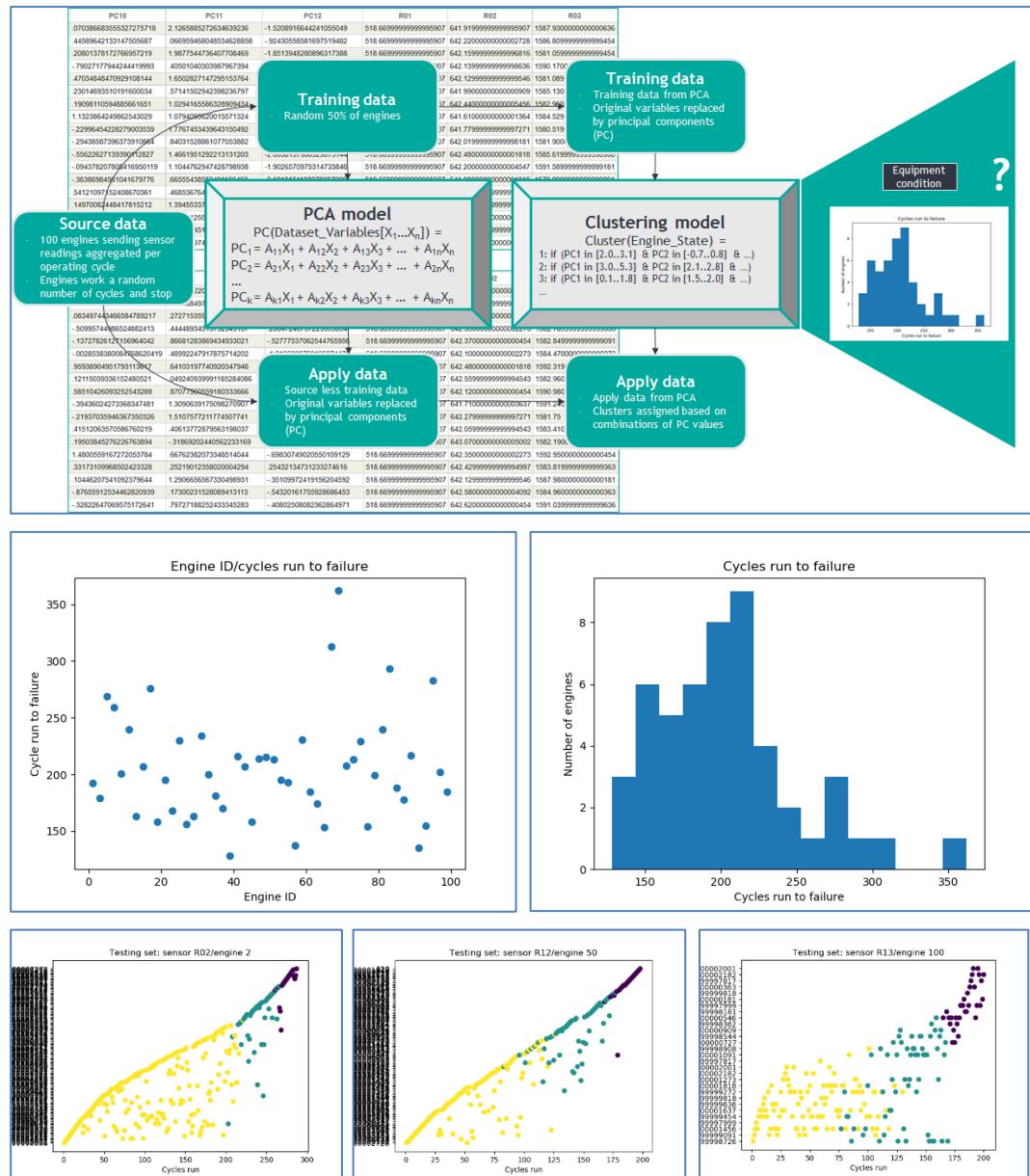
6.5. 002 Engine Condition Classification

Toolsets: ObjectScript, Python Gateway

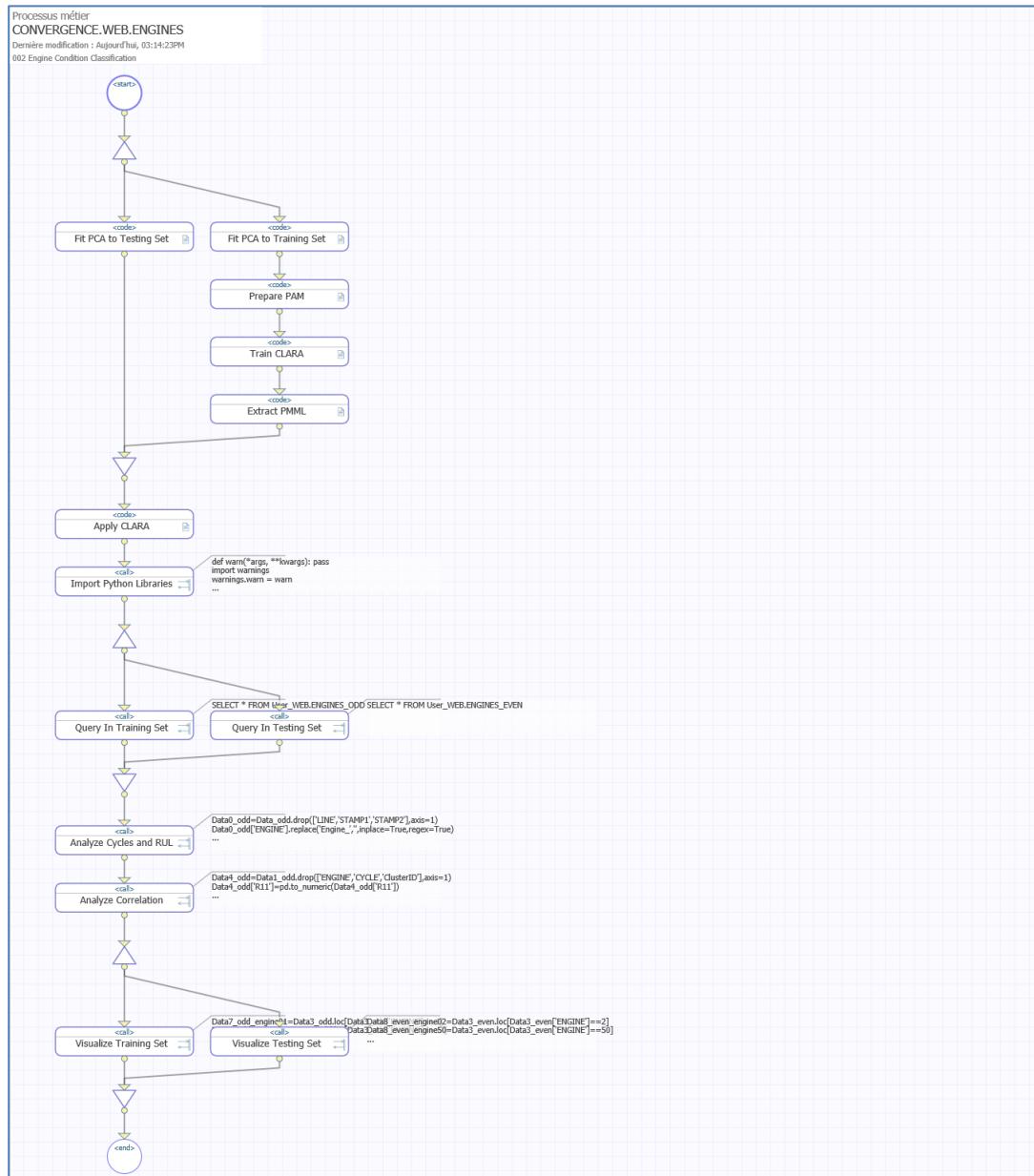
Algorithms: PCA, PAM, CLARA

6.5.1. Background

Object condition modeling: by applying dimensionality reduction and clustering methods to object state data, we dynamically determine presence of an object in “start”, “mid” or “end” phase of its lifecycle. Modeling results allow having a more efficient replacement process while avoiding unpredicted failures.



6.5.2. Implementation



6.5.3. Walkthrough

1. Fit PCA to Testing Set: calculates principal components representing the testing set of engines.
2. Fit PCA to Training Set: calculates principal components representing the training set of engines.
3. Prepare PAM: prepares the clustering model for being trained.
4. Train CLARA: trains the clustering model using a concrete algorithm.
5. Extract PMML: extracts clustering rules from the trained clustering model and converts them to PMML format.
6. Apply CLARA: applies the clustering model to the testing set to predict cluster numbers.
7. Import Python Libraries: loads the required libraries to Python context.

8. Query In Training Set: loads training data. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
9. Query In Testing Set: loads testing data.
10. Analyze Cycles and RUL: calculate additional metrics needed for further visual analysis of survived cycles and remaining useful life (RUL). **ACTION:** adjust the file paths to fit your environment.
11. Analyze Correlation: calculate correlation matrices to study dependencies among variables.
12. Visualize Training Set: generate several graphs to support visual analysis and validation of modeling results using training data.
13. Visualize Testing Set: generate several graphs to support visual analysis and verification of modeling results using testing data.

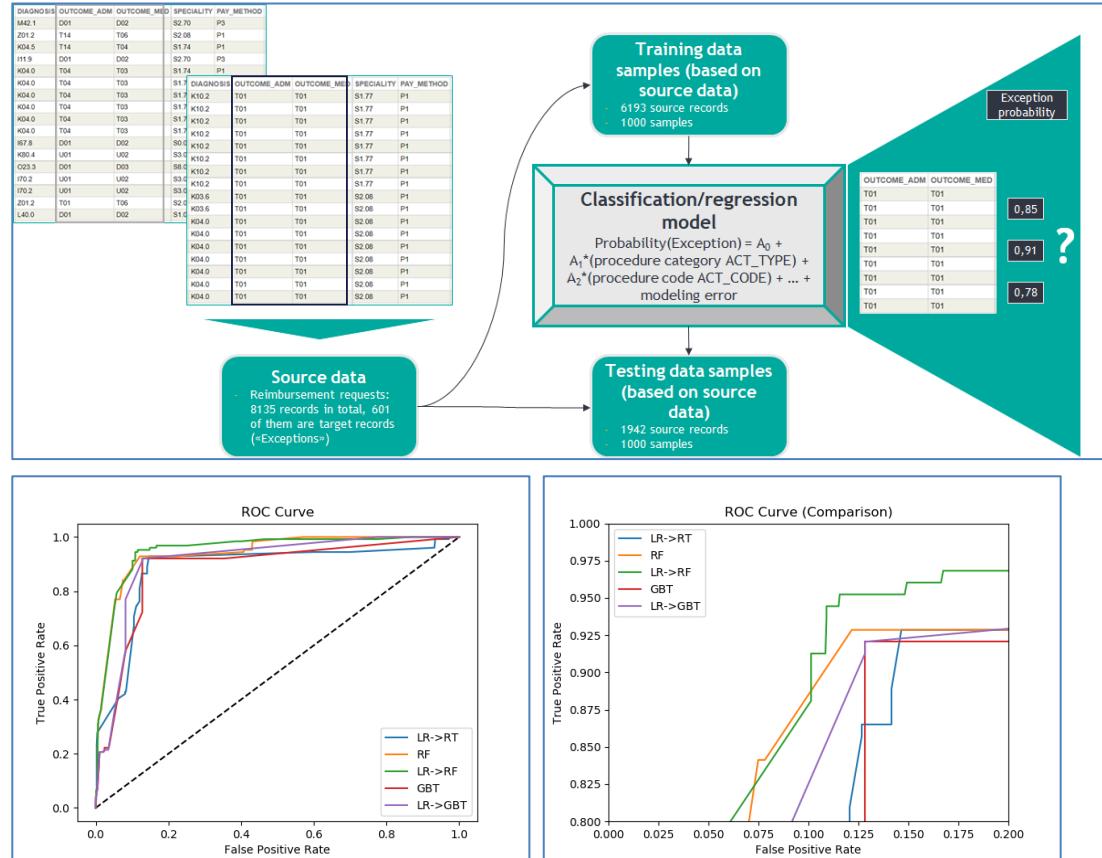
6.6. 003 Reimbursement Request Check

Toolsets: Python Gateway

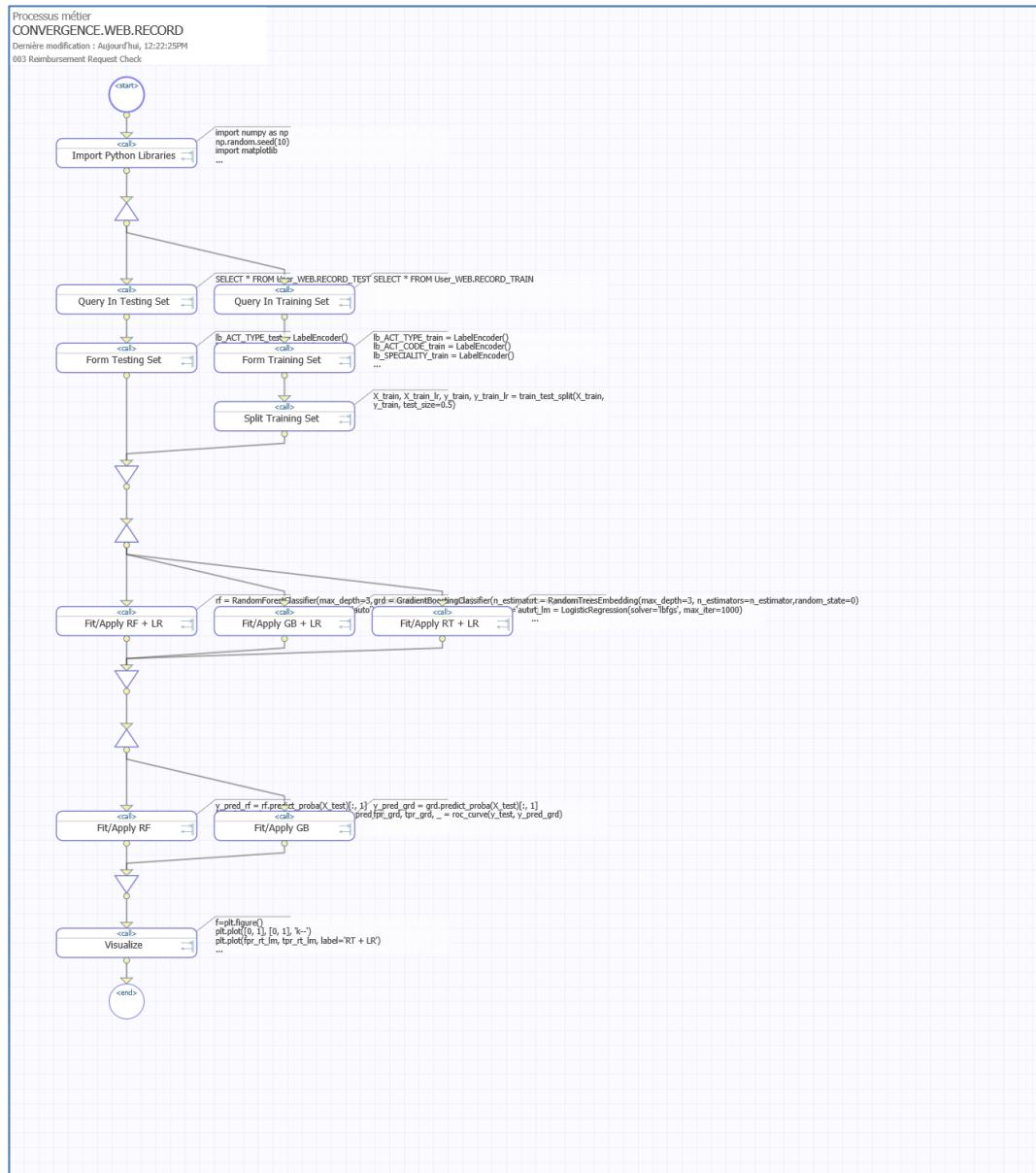
Algorithms: LR, RT, RF, GB, OneHotEncoder

6.6.1. Background

A classification model trained on samples of various types of deviations, detects them automatically in a flow of transactions.



6.6.2. Implementation



6.6.3. Walkthrough

1. Import Python Libraries: loads the required libraries to Python context.
2. Query In Testing Set: loads testing data. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
3. Form Testing Set: encodes testing data from text labels to numbers
4. Query In Training Set: loads training data
5. Form Training Set: encodes training data from text labels to numbers
6. Split Training Set: splits training data into two subsets – one to train linear regression, another one to train the other models in the showcase
7. Fit/Apply RT + LR: fits and applies a bundle of random tree and linear regression models (evaluation and testing modes)
8. Fit/Apply RF + LR: fits and applies a bundle of random forest and linear regression models (evaluation and testing modes)

9. Fit/Apply GB + LR: fits and applies a bundle of gradient boosting and linear regression models (evaluation and testing modes)
10. Fit/Apply RF: fits and applies a random forest model (evaluation and testing modes)
11. Fit/Apply GB: fits and applies a gradient boosting model (evaluation and testing modes)
12. Visualize ROC Curves: saves to graphical files the ROC charts (makes possible defining the winner model). **ACTION:** adjust the file paths to fit your environment.

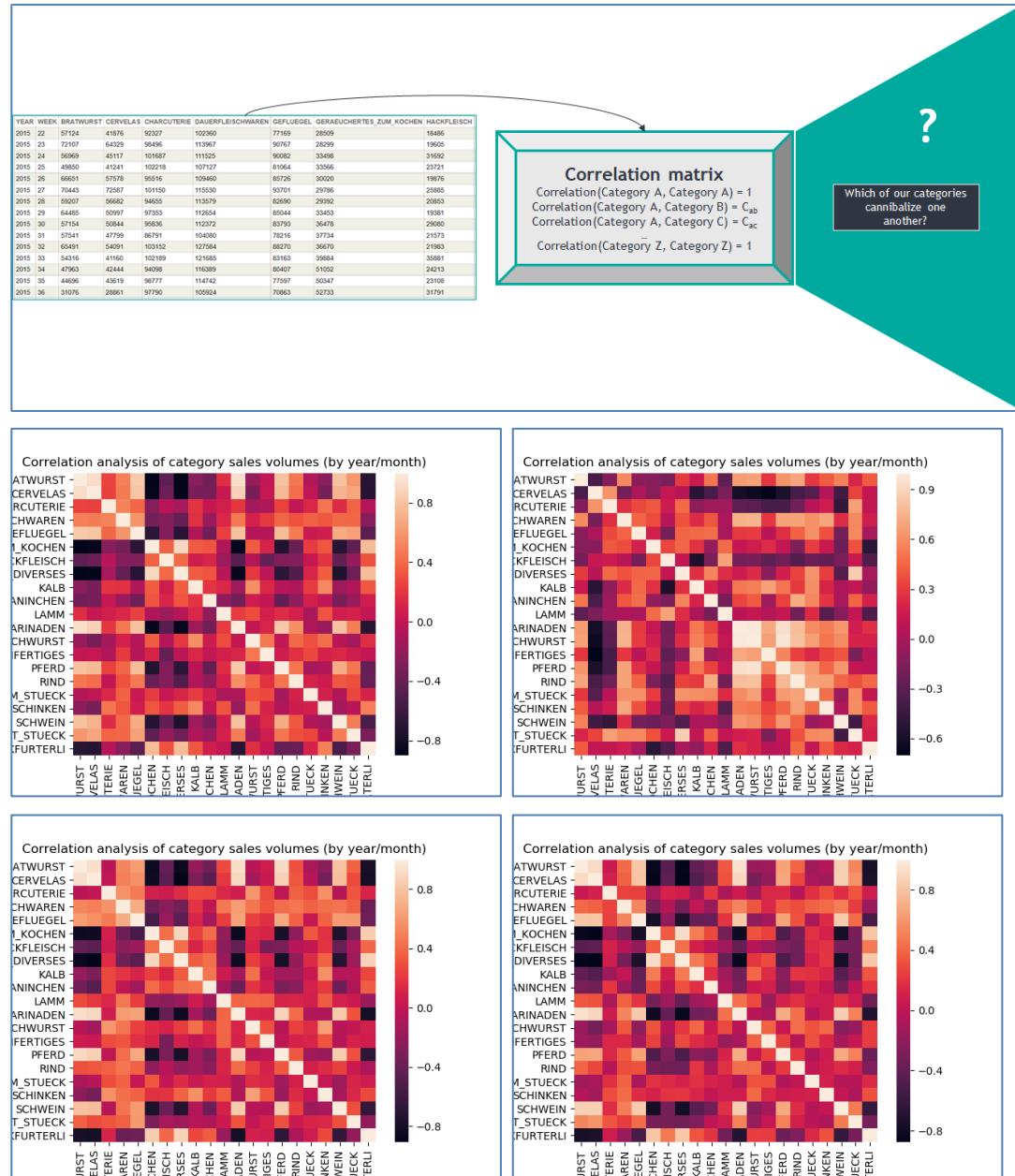
6.7. 004 Retail Cannibalization Analysis

Toolsets: Python Gateway

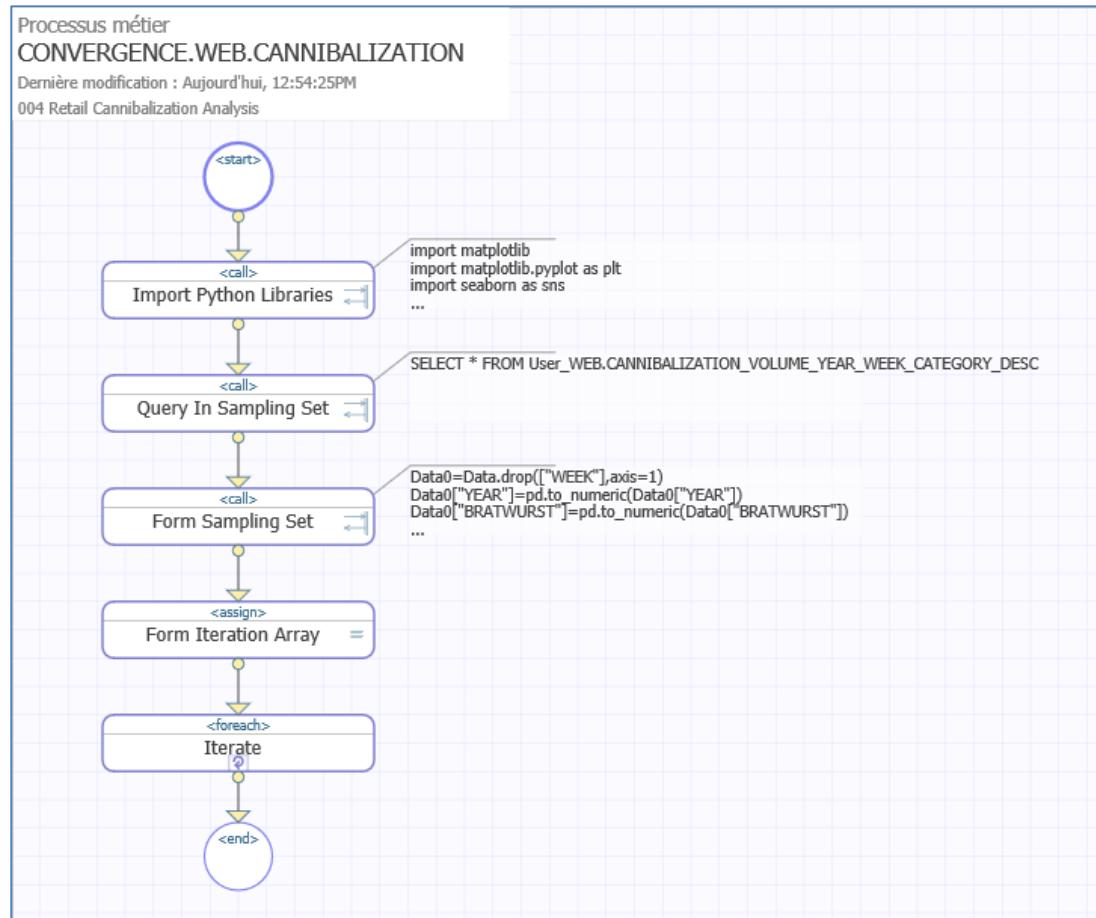
Algorithms: Statistical Functions

6.7.1. Background

Correlation analysis is a universal method to discover potential mutual influence among a set of variables. One of the examples is product cannibalization in retail: certain product categories lose in sales volume if their “cannibalizing” counterparts show growing sales volume.



6.7.2. Implementation



6.7.3. Walkthrough

1. Import Python Libraries: loads the required libraries to Python context.
2. Query In Sampling Set: loads sampling data. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
3. Form Sampling Set: prepares sampling data for a correlation analysis
4. Form Iteration Array: forms an array to iterate over while doing the correlation analysis (next step)
5. Iterate: iterates over the sampling data and builds sample correlation matrices per each iteration

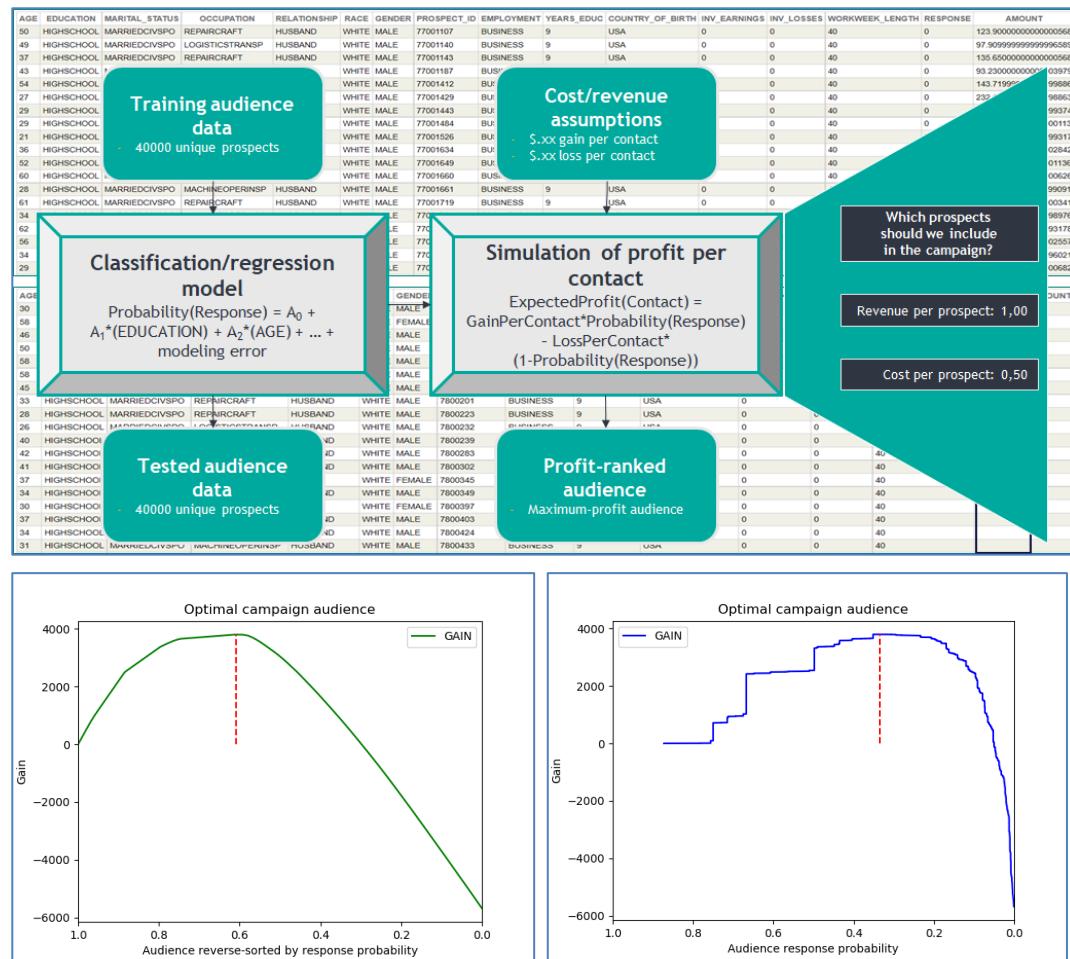
6.8. 005 Marketing Campaign Optimization

Toolsets: Python Gateway

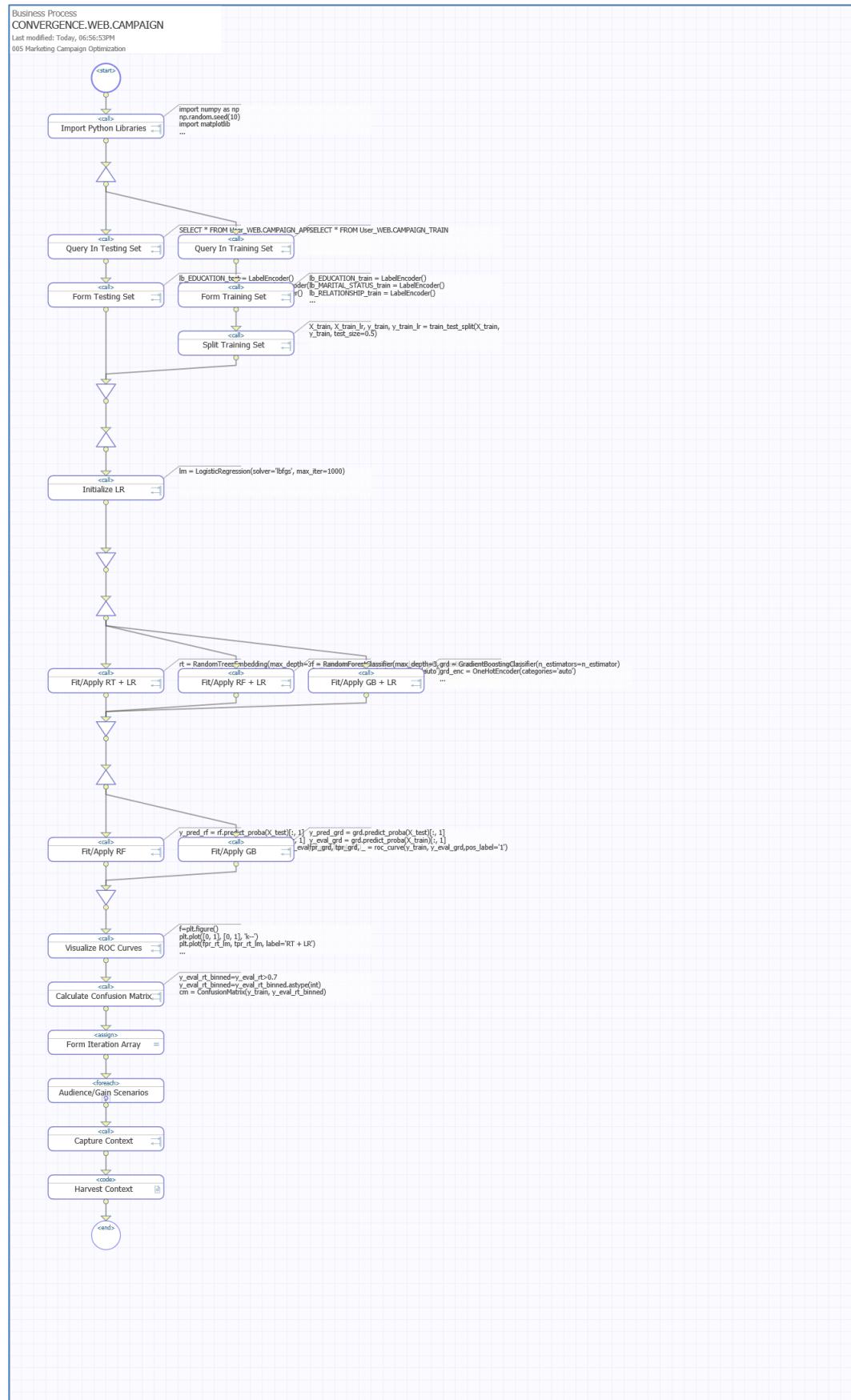
Algorithms: LR, RT, RF, GB, OneHotEncoder

6.8.1. Background

We implement a stack of classification models to have a robust estimate of response probability. We then use response probabilities to calculate the expected gain per prospect. We finally reverse-sort the audience on response probability and estimate the maximum summary gain for the campaign, as well as the part of the audience that generates it.



6.8.2. Implementation



6.8.3. Walkthrough

1. Import Python Libraries: loads the required libraries to Python context.
2. Query In Testing Set: loads testing data. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
3. Form Testing Set: encodes testing data from text labels to numbers
4. Query In Training Set: loads training data
5. Form Training Set: encodes training data from text labels to numbers
6. Split Training Set: splits training data into two subsets – one to train linear regression, another one to train the other models in the showcase
7. Initialize LR: initializes a linear regression model
8. Fit/Apply RT + LR: fits and applies a bundle of random tree and linear regression models (evaluation and validation modes)
9. Fit/Apply RF + LR: fits and applies a bundle of random forest and linear regression models (evaluation and validation modes)
10. Fit/Apply GB + LR: fits and applies a bundle of gradient boosting and linear regression models (evaluation and validation modes)
11. Fit/Apply RF: fits and applies a random forest model (evaluation and validation modes)
12. Fit/Apply GB: fits and applies a gradient boosting model (evaluation and validation modes)
13. Visualize ROC Curves: saves to graphical files the ROC charts (makes possible defining the winner model). **ACTION:** adjust the file paths to fit your environment.
14. Calculate Confusion Matrix: calculates a confusion matrix for the winner model
15. Form Iteration Array: forms an array to iterate over while doing the audience/gain sensitivity analysis (next step)
16. Audience/Gain Scenarios: iterates over a set of experimental scenarios to discover the range of possible optimal audience and maximum gain
17. Capture Context: saves all the objects currently existing in Python context to globals in IRIS
18. Harvest Context: extracts from the context saved in the previous step the confusion matrix and saves it as a separate global in IRIS with each node corresponding to a cell value in the confusion matrix.

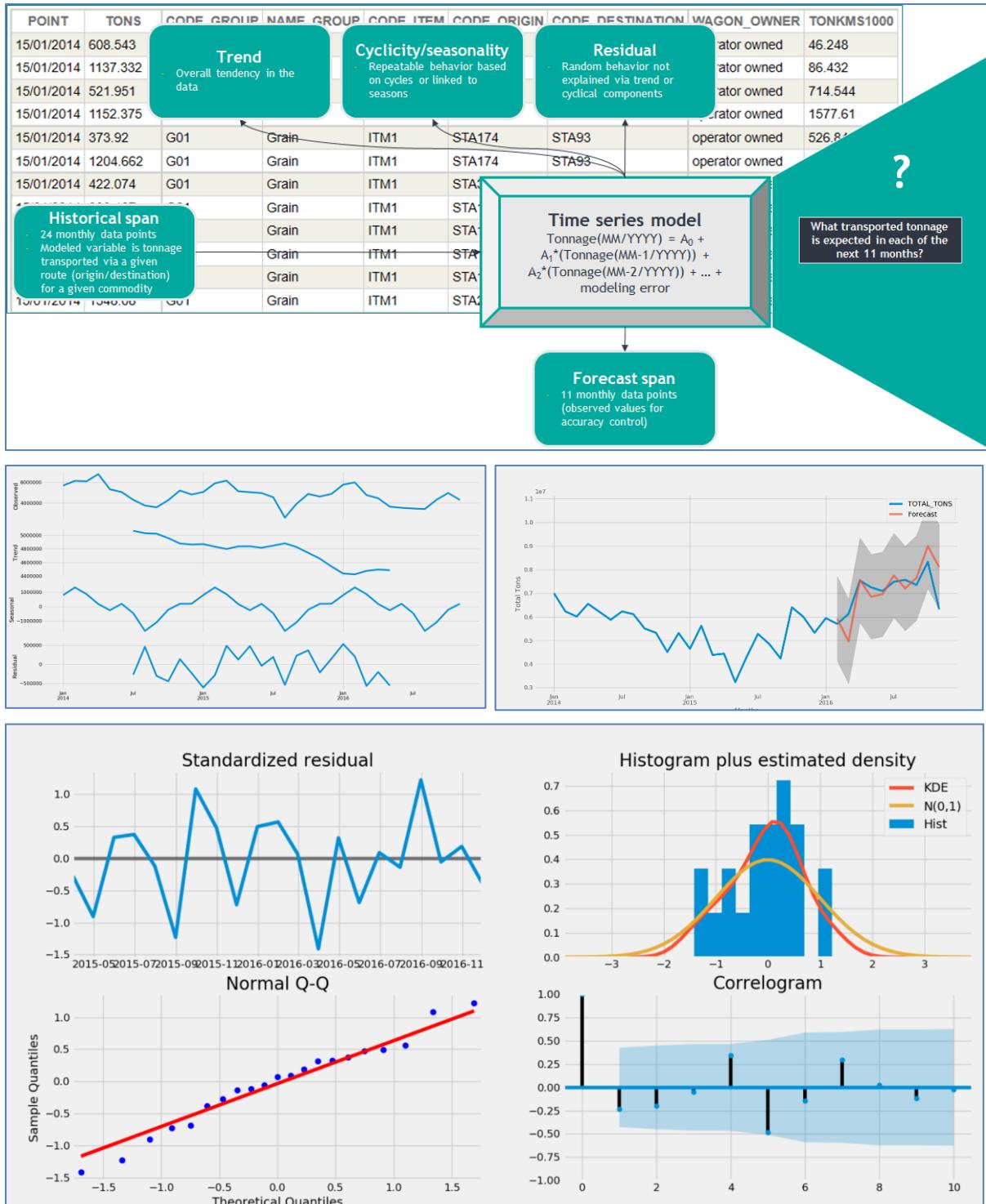
6.9. 006 Rail Time Series Discovery

Toolsets: Python Gateway

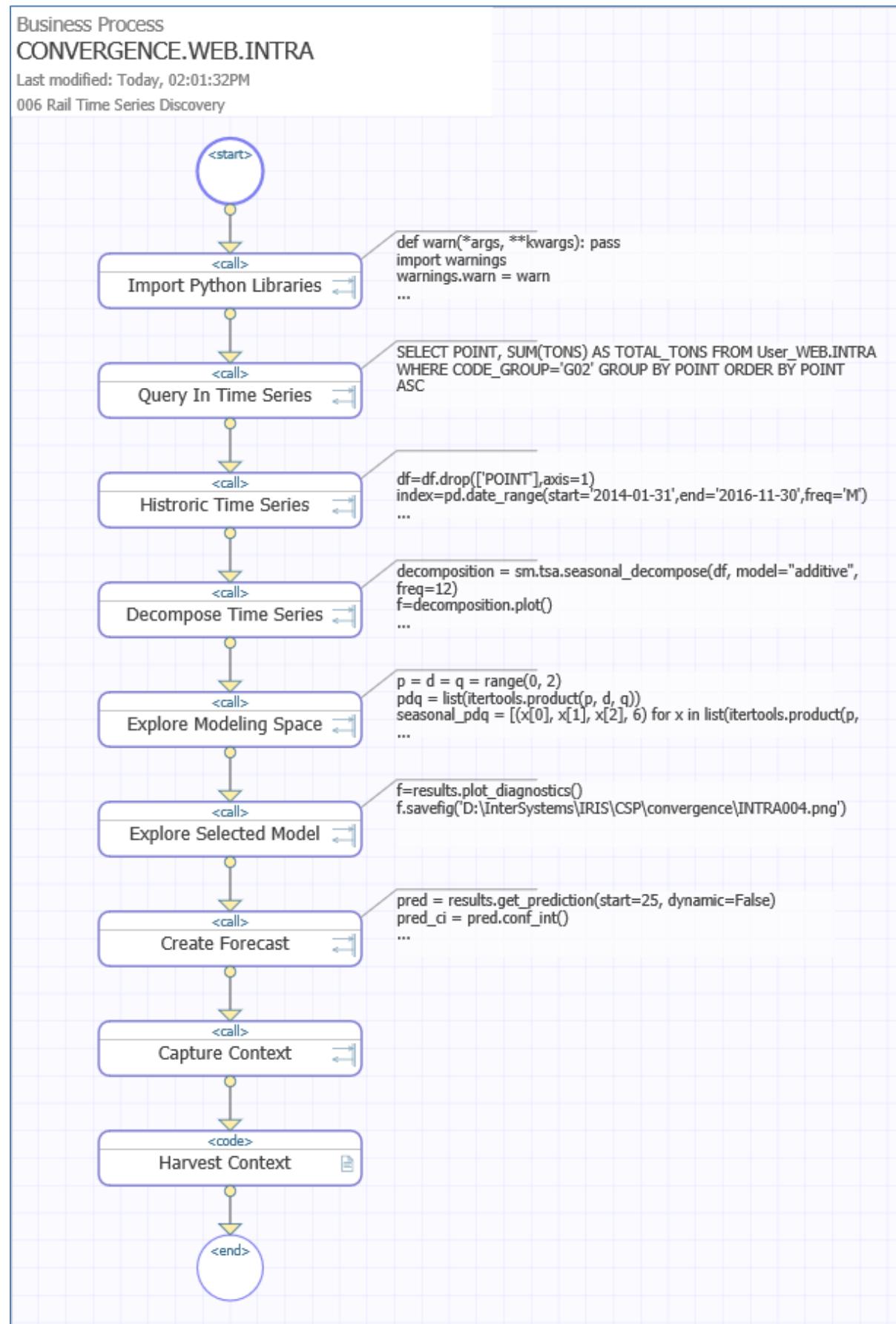
Algorithms: SARIMAX

6.9.1. Background

Time series analysis: regression-based methods of time series analysis allow a decomposition of the "signal" (historical values available at the moment of modeling) into several components: trend, cyclical and "noise" parts. Each component is modeled individually, modeling results added up form a forecast.



6.9.2. Implementation



6.9.3. Walkthrough

1. Import Python Libraries: loads the required libraries to Python context.
2. Query In Time Series: loads time series data. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
3. Historic Time Series: transform for the analysis and visualize historic time series data. **ACTION:** adjust the file paths to fit your environment.
4. Decompose Time Series: “split” the signal represented by time series data into trend, cyclical and noise components. **ACTION:** adjust the file paths to fit your environment.
5. Explore Modeling Space: iterate through various combinations of time series model parameters to study their impact on modeling accuracy. **ACTION:** adjust the file paths to fit your environment.
6. Explore Selected Model: calculate and visualize the various model metrics helping to interpret modeling results. **ACTION:** adjust the file paths to fit your environment.
7. Create Forecast: calculate and visualize a forecast using the trained model. **ACTION:** adjust the file paths to fit your environment.
8. Capture Context: saves all the objects currently existing in Python context to globals in IRIS
9. Harvest Context: extracts from the context saved in the previous step the time series dataframe and saves it as a separate global in IRIS.

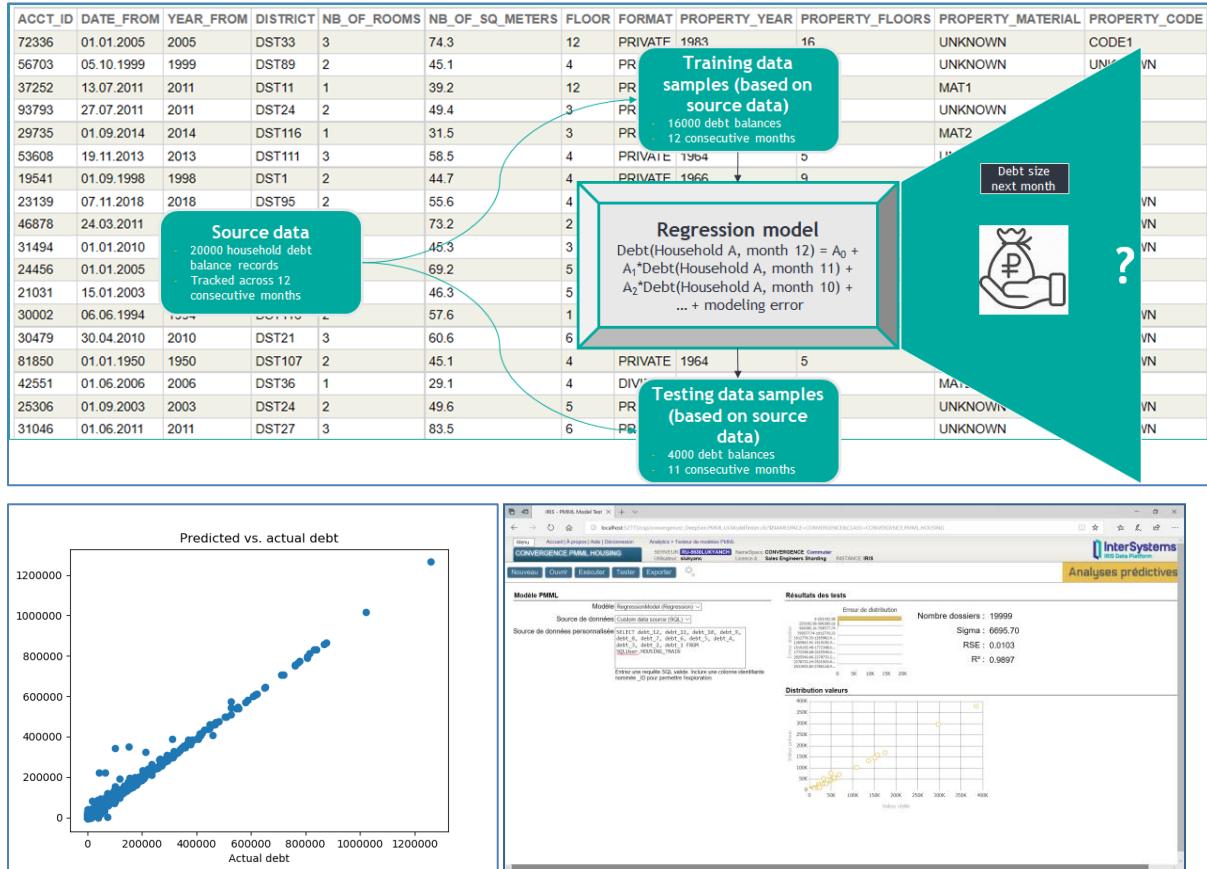
6.10. 007 Housing Debts Prediction

Toolsets: Python Gateway

Algorithms: LR

6.10.1. Background

A regression model trained on a history of debt behaviors is applied to a new dataset to obtain a prediction consisted with the observed debt numbers.



6.10.2. Implementation



6.10.3. Walkthrough

1. Import Python Libraries: loads the required libraries to Python context.
2. Query In Training Set: loads training data. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
3. Fit LR Model: fits a linear regression model.
4. Apply LR Model: applies a linear regression model.
5. Visualize Prediction Quality: saves to a graphical file the actual/predicted chart.
ACTION: adjust the file paths to fit your environment.
6. Harvest PMML: extracts clustering rules from the trained regression model and converts them to PMML format.

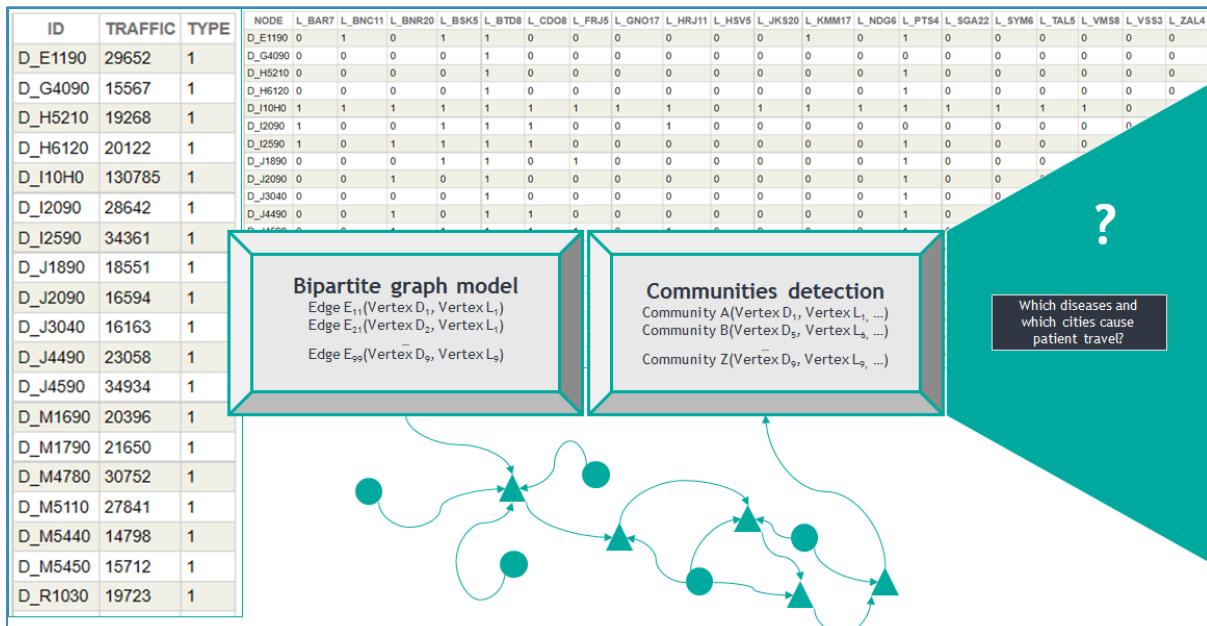
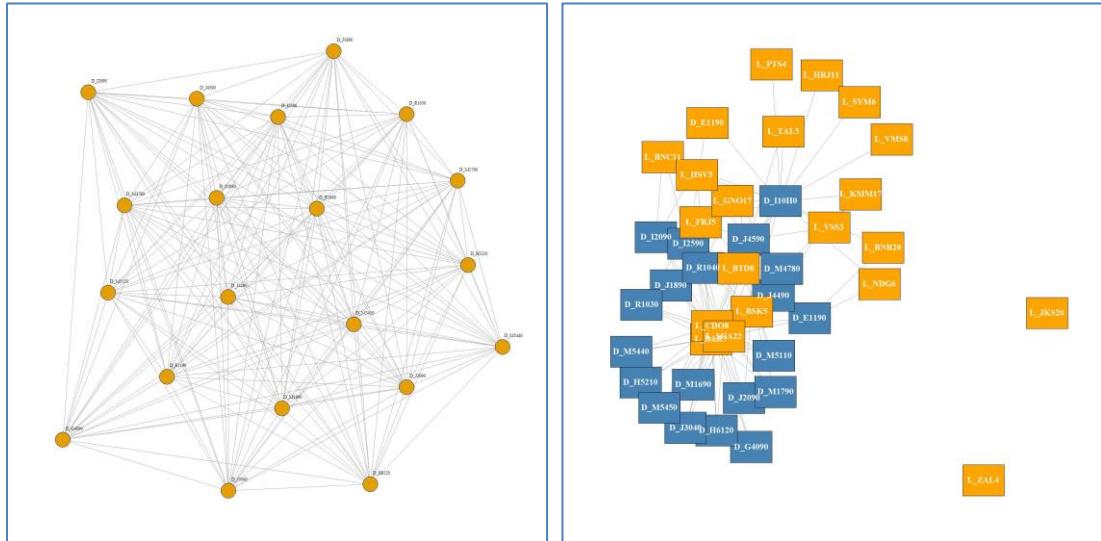
6.11. 008 Diseases Network Analysis

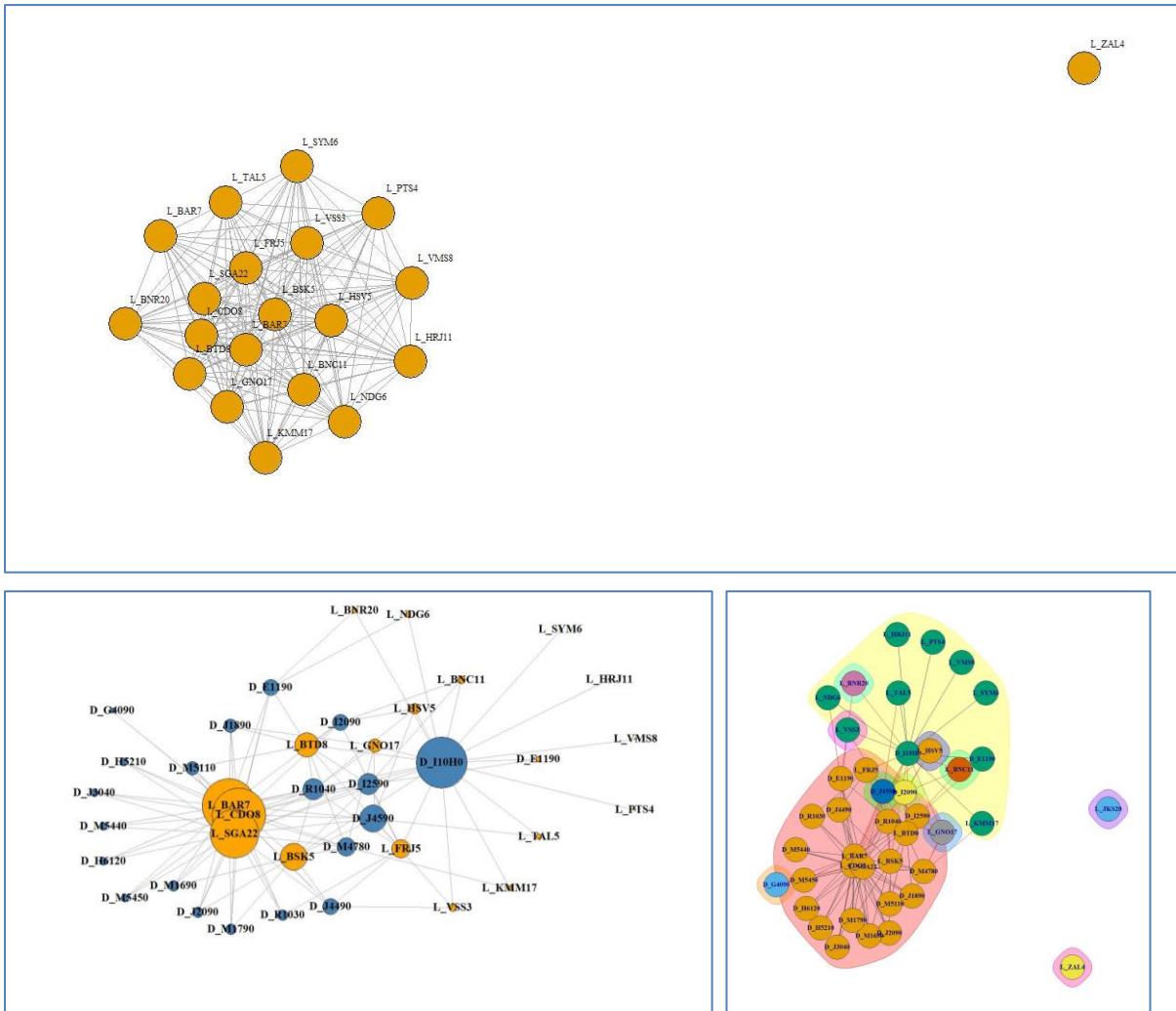
Toolsets: R Gateway

Algorithms: IGraph

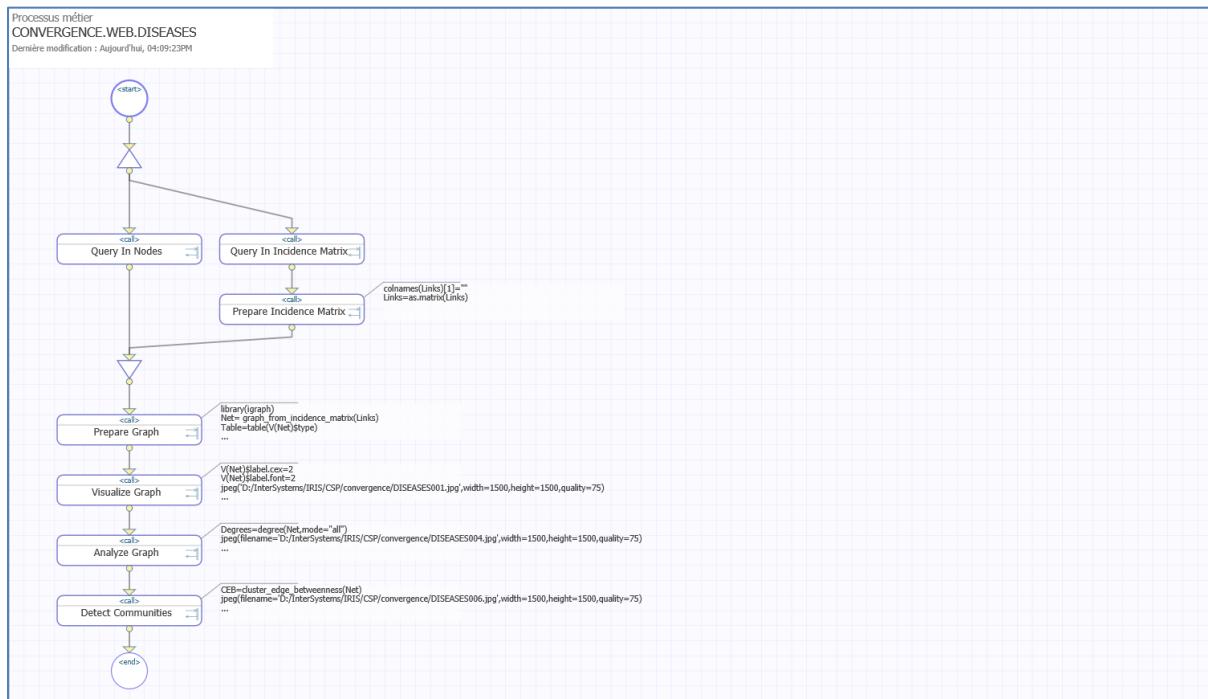
6.11.1. Background

A graph model is instrumental to study data dependencies that are easier to detect via a network rather than a table representation.





6.11.2. Implementation



6.11.3. Walkthrough

1. Query In Nodes: loads the graph nodes. **ACTION:** implement the data tables in the namespace(s) that fit your environment.
2. Query In Incidence Matrix: loads the graph incidence matrix.
3. Prepare Incidence Matrix: adjusts the incidence matrix dataframe for use in further graph analysis components.
4. Prepare Graph: defines and builds the graph.
5. Visualize Graph: saves to graphical files the graph projections. **ACTION:** adjust the file paths to fit your environment.
6. Analyze Graph: saves to graphical files the results of various graph analyses (bipartite visualization, links intensity, etc.).
7. Detect Communities: determines in the graph subsets of highly linked components, saves to a graphical file the communities visualization.

6.12. 009 Bell State Automation

Toolsets: Python Gateway

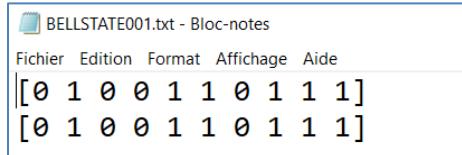
Algorithms: pyQuil

Idea: <http://docs.rigetti.com/en/latest/start.html#installation-and-getting-started>

ACTION: install and start `qvm` and `quilc` servers as described under the link above

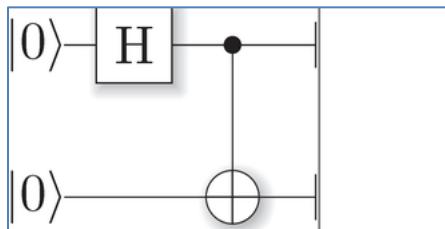
6.12.1. Background

A Bell state in quantum computations is a form of quantum entanglement under which the two entangled qubits show perfectly correlated measurements:



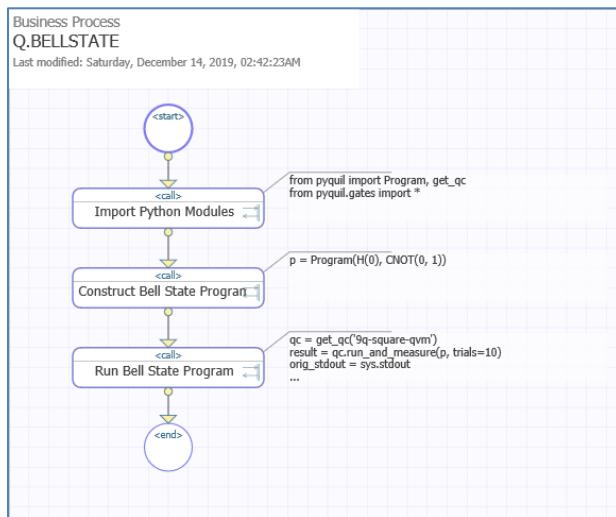
```
BELLSTATE001.txt - Bloc-notes
Fichier Edition Format Affichage Aide
[[0 1 0 0 1 1 0 1 1 1]
 [0 1 0 0 1 1 0 1 1 1]]
```

The most straightforward quantum circuit that implements a Bell state is like that:



In our showcase, we implement the above quantum circuit and run it 10 times.

6.12.2. Implementation



6.12.3. Walkthrough

1. Import Python Modules: loads the required modules to Python context.
2. Construct Bell State Program: constructs the quantum circuit object.
3. Run Bell State Program: runs calculations using the quantum circuit.



InterSystems Corporation
World Headquarters

One Memorial Drive
Cambridge, MA 02142-1356
Tel: +1.617.621.0600

InterSystems.com