Pick 3D Descriptors

Shaillay Kumar Dogra
Scientific Editor – QSAR World
editor@qsarworld.com

Notes:

1. This Jython script works in Sarchitect Designer version 2.2
2. Learn about Sarchitect Designer – http://www.strandls.com/sarchitect/index.html
3. Get Sarchitect – http://www.strandls.com/sarchitect/freetrial.php

The actual script follows this discussion. It is also accessible directly from the webpage in .py format.

Discussion:

A small but useful script that allows one to pick 3D (conformational) descriptors from a set wherein 'all' descriptors have been computed. This is of help if, say, one wants to build models on 3D descriptors only, then compare against models obtained on the full set (or the 2D descriptors only). If comparable models can be obtained across different descriptor sets, then models obtained with 2D descriptors are preferable since they are intuitive to understand and require lesser computation (no need for structure optimization etc.). Or it could so happen that for complex endpoints 3D descriptors may prove the best for modeling.

Running this script will create a subset with only the 3D descriptors. Input is a superset of descriptors that also contains these 3D descriptors, in particular, "DPO1" and "Electro" as these are used in the script to mark the start and end of 3D descriptor columns. If these are not found, a message gets prompted to that effect and the subset containing 3D descriptors is not created.

Another sister script does the complementary. It picks the 2D descriptors and creates a subset.

Cite this as:

Dogra, Shaillay K., "Script for picking 3D descriptors" from QSARWorld – free online resource for QSAR modeling.
http://www.qsarworld.com/virtual-workshop.php

```
##
##
## sarchitect designer 2.2 script to pick 3-D descriptors from a
## dataset and carry them into a child-datased labelled as "3D-
## descriptors"
##
##
## Author: Shaillay Kumar Dogra
## Date: July 06, 2007
## editor@qsarworld.com
##
##
## Note - "Marked" columns are carried into the child-dataset.
##        This script doesn't control that.
##
##
## There could be various ways of picking descriptors:
##
## a) based on names; will have to type out all the 3-D descriptor
## names and handle exceptions if some not found (some subset not
## computed by ticking it off)
##
## b) look for first 3D descriptor ("DP01"), find its index; find index
## of last 3D descriptor ("Electro") and thus get handle of all 3D
## descriptor indices assuming them to be in b/w these two - "DP01" &
## "Electro"
##
##
## Implementing latter option-b in this script
##
## Will give an error message if start-point ('DP01') and/or stop-point
## ('Electro') is not found.
## Will not create a child dataset in such cases and give a message to
## that effect.
##
##
##


import script
from script.dataset import *
from script.algorithm import *
from script.project import *
from script.omega import createComponent, showDialog
from javax.swing import *
from math import *

node = getActiveProject().getActiveDatasetNode()
dataset = node.getDataset()

startcol = dataset.getColumn('DP01')
startindex = dataset.index(startcol)
#print startcol, startindex
if (startindex == -1):   # Happens if 'DP01' not found
        parent=script.tool.getTool().getFrame()
        mesg = "Start point 'DP01' not found!"
```

```
        JOptionPane.showMessageDialog(parent,mesg,"STATUS!",JOptionPane.INFORMATIO
N_MESSAGE)


stopcol = dataset.getColumn('Electro')
stopindex = dataset.index(stopcol)
#print stopcol, stopindex
if (stopindex == -1):    # Happens if 'Electro' not found
        parent=script.tool.getTool().getFrame()
        mesg = "Stop point 'Electro' not found!"
        JOptionPane.showMessageDialog(parent,mesg,"STATUS!",JOptionPane.INFORMATIO
N_MESSAGE)


colindex = [ ]

i = startindex
if (startindex != -1 and stopindex != -1): # Happens if start-point "DP01" and stop-point "Electro"
was not found
        while i < (stopindex+1):
                colindex.append(i)
                i = i + 1


rowIndices=[i for i in range(dataset.getRowCount())]
colIndices= colindex
if (startindex != -1 and stopindex != -1): # Happens if start-point "DP01" and stop-point "Electro"
was not found
        node.addChildDatasetNode("3D-Descriptors", rowIndices, colIndices)
        script.view.Table().show()
else:
        parent=script.tool.getTool().getFrame()
        mesg = "No child-dataset created!"
        JOptionPane.showMessageDialog(parent,mesg,"STATUS!",JOptionPane.INFORMATIO
N_MESSAGE)


##
## END
##
```

End of Document