Help for using experimentation workflows

Before starting experimentation workflows open *Iterate through datasets (files).rmp* file located in Experimental workflows/Dataset Iteration directory using some text editor.

You should see something like Figure 1.


```
1  <?xml version="1.0" encoding="UTF-8" standalone="no"?>
2  <process version="5.1.008">
3    <context>
4      <input/>
5      <output/>
6      <macros/>
7    </context>
8    <operator activated="true" class="process" compatibility="5.1.008" expanded="true" name="Process">
9      <parameter key="logverbosity" value="init"/>
10     <parameter key="random_seed" value="2001"/>
11     <parameter key="send_mail" value="never"/>
12     <parameter key="notification_email" value=""/>
13     <parameter key="process_duration_for_mail" value="30"/>
14     <parameter key="encoding" value="SYSTEM"/>
15     <process expanded="true" height="462" width="768">
16       <operator activated="true" class="loop_files" compatibility="5.1.008" expanded="true" height="76" name="Loop Files" width="90" x="45
17         <parameter key="directory" value="C:\eclipse\workspace\RapidMiner_Vega\\Experimentation workflows\Data examples"/>
18         <parameter key="filter" value="train_.*.aml"/>
19         <parameter key="file_name_macro" value="file_name"/>
20         <parameter key="file_path_macro" value="file_path"/>
21         <parameter key="parent_path_macro" value="parent_path"/>
22         <parameter key="recursive" value="false"/>
23         <parameter key="iterate_over_files" value="true"/>
24         <parameter key="iterate_over_subdirs" value="false"/>
25         <process expanded="true" height="462" width="768">
26           <operator activated="true" class="generate_macro" compatibility="5.1.008" expanded="true" height="76" name="Generate Macro (2)"
27             <list key="function_descriptions">
28               <parameter key="train_file" value="&quot;C:\\eclipse\\workspace\\RapidMiner_Vega\\Experimentation workflows\\Data examples\\
29               <parameter key="test_file" value="&quot;C:\\eclipse\\workspace\\RapidMiner_Vega\\Experimentation workflows\\Data examples\\
30             </list>
31             <parameter key="use_standard_constants" value="true"/>
```

Figure 1. Iterate through datasets (files).rmp

Change lines 17, 28 and 29 which contain paths to your data examples and put local paths on your computer.

Now you can run RapidMiner and open Iterate through datasets (files) workflow.

In LoopFiles you can choose which sub process you would like to run and what data sets from your data set folder (Figure 3). Sub processes for all operators are saved in folder Time Measurement workflows.
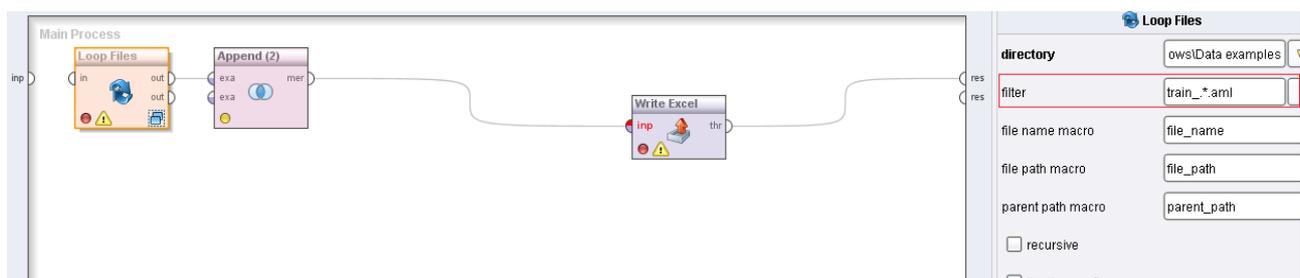
Figure 2. Iterate through datasets (files) workflow

Use filter to specify regular expression to filter the names of files wanted in training set. Note that your test files should have identical names as train files just with extension test.
For example if you have train5.aml train file then your corresponding test file should be test5.aml.
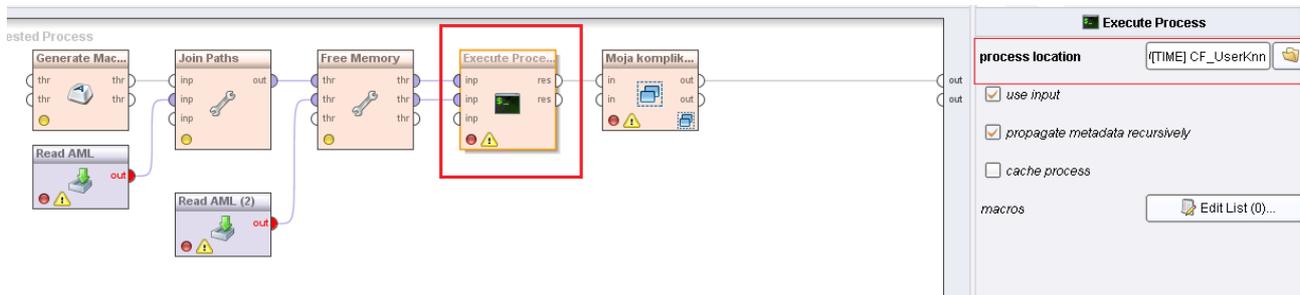
Figure 3. Defining which operator to test using Execute Process (Loop files operator)

If you would like to measure execution time of some other operator you should specify it in Log operator inside Loop operator of Time measurement workflow.
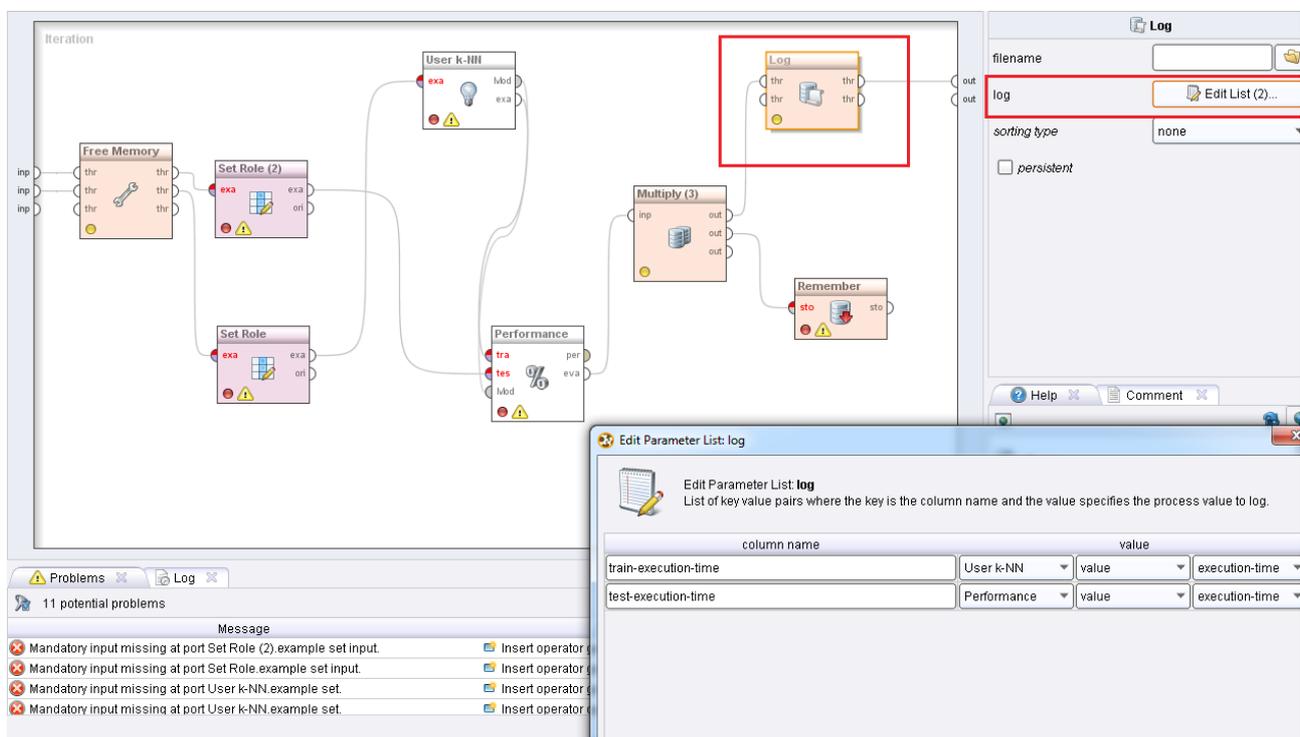
Figure 4. Defining information to measure execution time (Loop operator in Time measurement sub process)