Integration of AutoML and ARX

1. Development Environment
   1. The software is modified based on ARX version 3.4.1 (Download at <http://arx.deidentifier.org/?ddownload=1922>) .
   2. The software is tested under Java(TM) SE Runtime Environment (build 1.8.0\_92-b1) (Download at <http://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html> ).
   3. The software is developed under Java Development Kit 1.8.0\_92 (Download at <http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html> ).
2. Changes to be added to AutoML
   1. Changes to AutoML UI

|  |  |
| --- | --- |
| **Path:** | (new) [path] the path for arx to store output files; hierarchy files for columns are also stored in this path |
| **Input File:** | [example.csv] Input csv file |
| **Model:** | Classification/regression/clustering |
| **Maximum Time per Model (seconds):** | e.g. 300 |

* 1. Call ARX from AutoML: find the part of code in AutoML that reads the input csv file, and add the following code:

**java –jar arx.jar [path] [example.csv]**

For example: **java arx.jar data data/example.csv**

(to have it work on Mac OS, use **java –XstartOnFirstThread –jar arx.jar [path] [example.csv]** instead)

* 1. Run ARX algorithms on files generated by AutoML: when ARX exits, read file **[path]/output.txt** in AutoML, pseudocode:

**file = open\_file([path]/output.txt)**

**for each line in file**

**input\_data = open\_file(line)**

**run AutoML algorithms on input\_data, and save results**

**end for**

Naming Conventions:

**output.txt** – Stores the filenames of the output data for the current program. This file is reset every time the user opens a new ARX program.

**0.csv, 1.csv, 2.csv** … – The output data files after anonymization. When the user exports a new output data file, a new file is stored, and the file name is added to **output.txt**.

**0\_config.txt, 1\_config.txt, 2\_config.txt** … -- The privacy configurations of the corresponding output files.

1. Changes added to ARX
   1. Class Name: **Main.java** (org.deientifier.arx.gui)
      1. Method Name: **main()** Status: Changed

* If there are more than two command line arguments, and the second one ends with “.csv”, call the loadFile() method to create a new project and load the input file.
* If the argument path doesn’t end with “\” or “/”, add “/” “\”
  1. Class Name: **MainWindow.java** (org.deientifier.arx.gui.view.impl)
     1. New Attributes:
* **String outPath** – the path to store output files
* **String outSpec** – the path to the output specification file “output.txt”
* **int savedCnt** – Count of saved output files
  + 1. Method Name: **setOutPath(String outPath)** Status: New Method
* Store the [path] given by AutoML into the outPath variable
* Create the “output.txt” file under [path]
  + 1. Method Name: **getOutPath()** Status: New Method
* Return the outPath variable
  + 1. Method Name: **getSavedCnt()** Status: New Method
* Get the count of saved output files
  + 1. Method Name: **increaseSavedCnt()** Status: New Method
* Increase the count of saved output files by one
  + 1. Method Name: **getOutSpec()** Status: New Method
* Return the path to the output specification file "output.txt"
  + 1. Method Name: **getMenuFile()** Status: Changed
* Added new item in Menu: “File – Export to AutoML”
* The button only shows when the model has output, the window under the tab “analysis”, and the outPath variable is not empty (“”)
  + 1. Method Name: **getHelpMenu()** Status: Changed
* Added new item in help menu: “Help – AutoML Instructions”
  + 1. Mehtod Name: **showAutoMLDialog()** Status: New Method
* Create a dialogue for AutoML instructions
  1. Class Name: **Controller.java** (org.deientifier.arx.gui)
     1. Method Name: **actionMenuFileExit()** Status: Changed
* Changed the choice items in the exit pop up window
* Added new choice: “save to AutoML and exit”
* When the user click save, save the output files to AutoML
  + 1. Method Name: **actionMenuFileExportDataToAutoML()** Status: New Method
* The method called when user clicks the button “Export to AutoML”
* Create a new WorkerAutoMLSaver class to save the outputs and privacy configurations
  + 1. Method Name: **increaseSavedCnt()** Status: New Method
* Increase the savedCnt variable in class MainWindow.java
  + 1. Method Name: **actionMenuFileLoad()** Status: New Method
* Create a new project
* Load the [example.csv] file into the new project
* If there is any file with the name [column\_name.csv] under [path] folder, load these files as hierarchy files, and set the corresponding columns to be quasi-identifying
* Pop out instructions on how to use AutoML
  + 1. Method Name: actionMenulHelpAutoML() Status: New Method
  1. Class Name: **WorkerAutoMLSaver.java** (New Class) (org.deientifier.arx.gui.worker)
* Save the privacy configurations and output files
  1. Class Name: **DialogAutoMLInfo.java** (New Class) (org.deidentifier.arx.gui.view.impl.menu)
* A dialogue for AutoML instructions (similar to DialogAbout.java class)
  1. Class Name: EditorCriterion.java (org.deidentifier.arx.gui.view.impl.menu)
     1. Method Name: crateInstructions() Status: New Method
* Create a label with explanations for the privacy model
  1. Class Names: **EditorCriterionDifferentialPrivacy.java, EditorCriterionDDisclosurePrivacy.java, EditorCriterionDPresence.java, EditorCriterionKAnonymity.java, EditorCriterionKMap.java, EditorCriterionLDiversity.java, EditorCriterionRiskBased.java, EditorCriterionTCloseness.java** (org.deidentifier.arx.gui.view.impl.menu)
     1. Method Name: build() Status: Changed
* Added one line to construct the instruction label
  1. File Name: **messages.properties** (org.deientifier.arx.gui.resources)
     1. New Messages: AutoML.1 – AutoML.13
* Different pop up messages used to link ARX and AutoML