# Description

This file describes the minor changes made by CEBS to the core TGx HDACi Classifier R code.

This information is accurate as of February 2023.

# Updates

## tgx\_classfier.R

* Updated the usage comments
  + To reference one of the new sets of example input files instead of the old one
  + To include the metadata/info file parameter
* Replaced all old example file references in comments with new version
* Line 31 (the *options(scipen-99)* line) was *uncommented* to restore scientific notation usage in the classifier. A comment was also added.
* References to “txt” file types were changed to “tsv” to reflect the actual expected format of the input files.
* References to “annotation” were changed to “metadata” because “annotation” has a different meaning in the scientific context.
* Bug fix: Removed line 145 “classified$Name <-- NULL” because this produced empty results.
* Bug fix: Lines 150-154 added to properly populate the results\_summary array which is being used on line 155.
* Bug fix: Line 129 needed to be changed so the name\_label for the unknown chemical is properly applied to most of the output files for the DDI classifer.
* References to “dose” and “dose\_unit” were changed to “concentration” and “concentration\_unit” respectively.

## code/classifier\_utils.R

* References to “DDi” were changed to “DDI”
* References to “txt” file types were changed to “tsv” to reflect the actual expected format of the input files.
* Restoring Scientific notation:
  + The **classifier\_hdaci** function:
    - Line 33 was *uncommented* to be active. (HDACi predictions, no rounding)
    - Line 34 was *commented out* to be inactive (HDACi predictions, rounding)
    - Line 35 was changed to use the already-computed **probNH** value from line 30 rather than recomputing the value again.
    - Comments were appended to lines 33-35 to describe what each line is used for.
    - Bug fix: Return value was changed to just return the **predictions** array.
  + The **classifier\_ddi** function:
    - Lines 176-181 configure scientific notation output. The original “rounding” lines are commented out and documented.
    - Bug fix: Return value was changed to just return the **predictions** array.

## code/common\_utils.R

* References to “dose” and “dose\_unit” were changed to “concentration” and “concentration\_unit” respectively.

## code/consts\_utils.R

* References to “DDi” were changed to “DDI”
* The **get\_txt\_file\_labels** function:
  + Lines 46 and 47 – changed “data” in labels to say “dist”, which seems to match better with the output.

## code/file\_io\_utils.R

* Lines 39-42 replace the original single “write.table” line to fix an output bug in the data files.
* Lines 52-56 replace the original single “write.table” line to fix an output bug in the data files.
* Lines 57, 69, 81:
  + Bug fix: Replaced **filename** in the function calls with **file\_out**.

## code/plot\_utils.R

* References to “DDi” were changed to “DDI”
* Renamed “cluster” file names to “dendro+pca” for clarity.

## examples/\*

* All example files were replaced with the blinded versions we are using in the new tool. The old example files were not all completely relevant and some of them had incorrect formatting.