

Evaluation

After you have performed the classification using TM Cleaner you would like to see how well the algorithm performed. For this you should annotate a sample from your test set, classify it and then compare the annotation with the automatic classification.

The script “EvaluateExamples.py” helps you evaluate the performance of the classifier.

Put the classified file and the manually annotated file in an empty directory. For the sake of this tutorial let’s assume that the directory is called Evaluation and:

- The manually annotated file is called: *about-small-en-it-annotated.txt*
- The classified file is called: *about-small-en-it-classified.txt*

You already know the format of the classified file. The format of the manually annotated file should be identical with the format of the training file:

0###0###source segment###target segment###category

1###1###source segment###target segment###category

To see an example of manually annotated and classified files look inside the directory “/Resources/Examples/Hunalign/Evaluation/En-It-Hunalign-LR”

Performing the evaluation

Run the program:

```
python EvaluateExamples.py --classified Evaluation/about-classified.txt --annotated Evaluation/about-manual.txt
```

The Output is in the directory Evaluation:

- statistics-sample.txt – contains the confusion matrix, precision, recall and F1 score for positive and negative classes and the balanced accuracy measure. The confusion matrix is constructed as follows:

		Predicted	
		1	0
	True		
	1	<i>True Positives</i>	<i>False Negatives</i>
	0	<i>False Positives</i>	<i>True Negatives</i>

- [miss-sample.txt](#) – contains the miss-classified lines.
- [visualize-sample.html](#) – is an html file you can load in your browser. On each line you see the source segment, the target segment the automatic category and the manual category for the sample. For the lines where the classifier made mistakes the cell at the intersection of the line and the automatic category column is colored in green and the content is read.