**Documentation for running the baseline systems – Group 04**

Step 1: running the models themselves:

**Task A monolingual:**

python3 baseline\_task\_A.py --train\_file\_path subtaskA\_train\_monolingual.jsonl --test\_file\_path subtaskA\_dev\_monolingual.jsonl --prediction\_file\_path subtaskA\_distilbert\_predictions\_monolingual.jsonl --subtask A --model 'distilbert-base-uncased'

**Task A multilingual:**

python3 baseline\_task\_A.py --train\_file\_path subtaskA\_train\_multilingual.jsonl --test\_file\_path subtaskA\_dev\_multilingual.jsonl --prediction\_file\_path subtaskA\_distilbert\_predictions\_multilingual.jsonl --subtask A --model 'distilbert-base-multilingual-cased'

**Task B:**

python3 subtaskB/baseline/transformer\_baseline.py --train\_file\_path subtaskB\_train.jsonl --test\_file\_path subtaskB\_dev.jsonl --prediction\_file\_path subtaskB\_distilbert\_predictions\_monolingual.jsonl --subtask B --model 'distilbert-base-uncased'

**Task C:**

**-**

This results in a predictions file as specified under the ‘--prediction\_file\_path’ argument. This prediction file can then be evaluated using scorer.py. This python file also uses format\_checker.py so make sure you also have that file installed in the same directory.

**Task A monolingual:**

python3 subtaskA/scorer/scorer.py --gold\_file\_path=subtaskA\_dev\_monolingual.jsonl --pred\_file\_path=subtaskA\_distilbert\_predictions\_monolingual.jsonl

**Task A multilingual**

python3 subtaskA/scorer/scorer.py --gold\_file\_path=subtaskA\_dev\_multilingual.jsonl --pred\_file\_path=subtaskA\_distilbert\_predictions\_multilingual.jsonl

**Task B:**

python3 subtaskB/scorer/scorer.py --gold\_file\_path=subtaskB\_dev.jsonl --pred\_file\_path=subtaskB\_distilbert\_predictions\_monolingual.jsonl

**Task C:**

**-**