Text-in-Context: Token-Level Error Detection for Table-to-Text Generation

Zdeněk Kasner
Simon Mille
Ondřej Dušek

kasner@ufal.mff.cuni.cz
simon.mille@upf.edu
odusek@ufal.mff.cuni.cz

Inputs
(i) Rotowire basketball data
(ii) data descriptions generated by neural NLG systems

Outputs
factual errors in the descriptions detected on token-level

How-to in three steps:

① Generate facts from the input table with a rule-based NLG system.

② Select the relevant subset of facts based on the similarity of sentence embeddings (the fact vs. the checked sentence).

③ Train RoBERTa to label the tokens with error categories based on the selected subset of facts (=context) on human-annotated data from organizers & synthetic from RotoWire train set.

Results: 1 out of 4 automatic metrics in the Shared Task.

| Error | Recall | Precision |
|-------|--------|-----------|
| NAME  | 0.750  | 0.846     |
| NUMBER| 0.777  | 0.750     |
| WORD  | 0.514  | 0.483     |
| CONTEXT| 0.000  | -         |
| NOT_CHK. | 0.000  | -         |
| OTHER | 0.000  | -         |
| Overall| 0.691  | 0.756     |

c = context size (# selected facts), EMR = synth data error level (% replaced entities)

Presented at INLG 2021, online.

More information at: https://github.com/kasnerz/accuracySharedTask_CUNI-UPF