



deepQuest-py: Large and Distilled Models for Quality Estimation

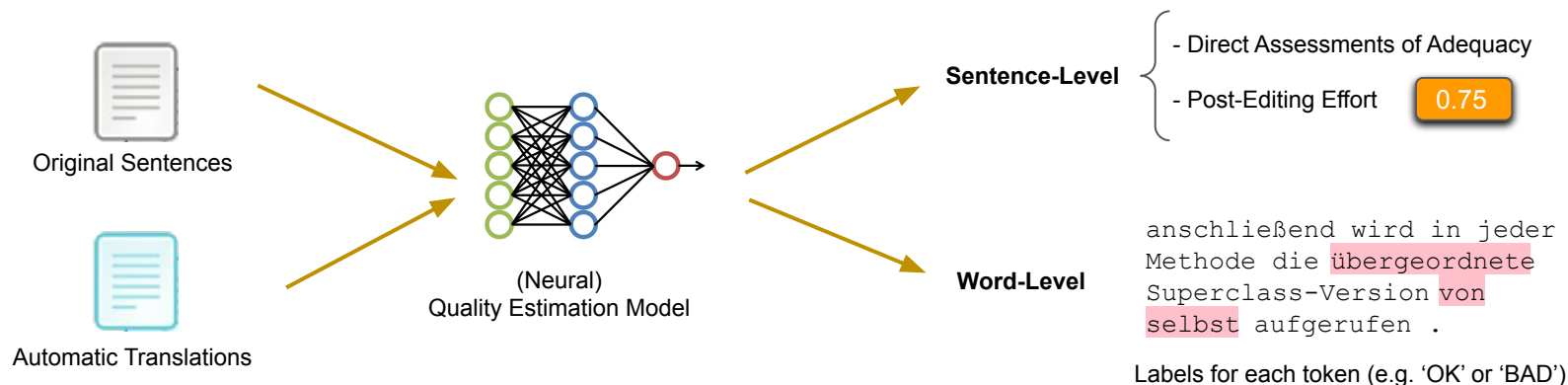
Fernando Alva-Manchego, Abiola Obamuyide, Amit Gajbhiye,
Frédéric Blain, Marina Fomicheva, Lucia Specia

EMNLP 2021 Demo

7-11 November 2021

Quality Estimation of Machine Translation

- To predict how good automatic translations are **without comparing them to gold-standard references**



Open-Source Software for Quality Estimation

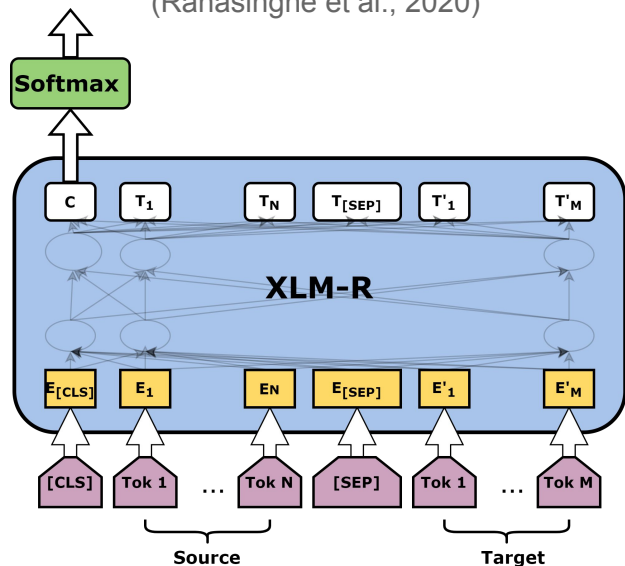
	 QuEst	 QuEst++	 deepQuest	 OpenKiwi	 TransQuest
Large SotA Architectures				✓	✓
Lightweight models	✓	✓	✓	✓	
Pre-trained QE Models					✓
Accessible for researchers	✓	✓	✓	✓	✓
Accessible for end-users (e.g. translators)				?	

Introducing **deepQuest-py**

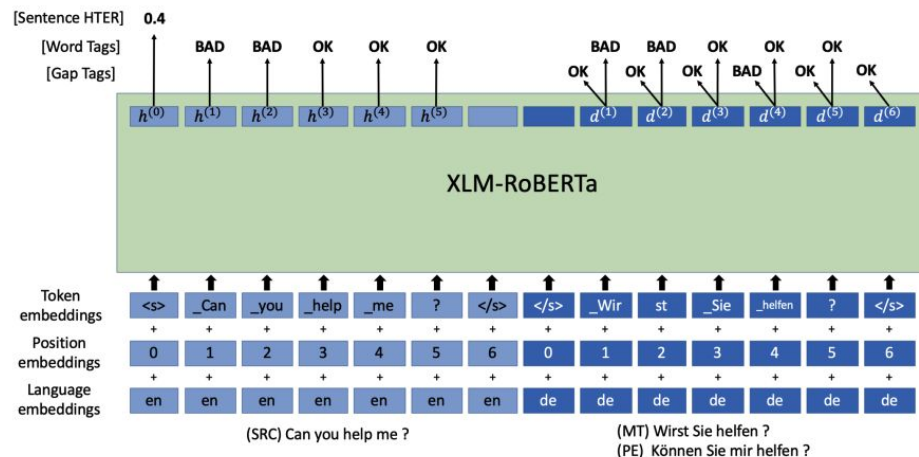
Large State-of-the-Art Models

Best models
of WMT20

Sentence-level (Ranasinghe et al., 2020)

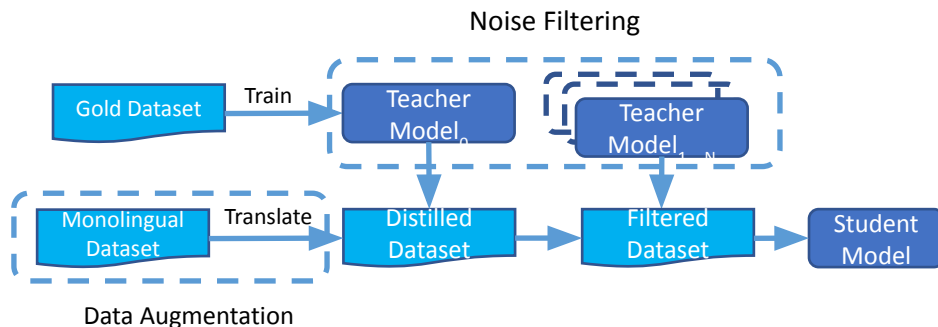


Word-level (Lee., 2020)



Light-Weight Distilled Models

Sentence-level (Gajbhiye et al., 2021)



Trade-off
performance-size

Smaller and
faster models!

Model	#params	Inference		
		Speed (secs.)	RAM (MiB)	Disk (M)
TransQuest _{XLM-R-Large}	561M	0.82	9,263.5	2,140
TransQuest _{DistillBERT}	135M	1.09	1,979.2	517
BiRNN	18M	0.39	155.6	132

Model	Language Pair				
	Et-En	Ro-En	Si-En	Ne-En	En-Zh
BiRNN _{Student+Aug}	0.50	0.69	0.45	0.54	0.17
TransQuest _{DistillBERT}	0.62	0.78	0.51	0.61	0.36

Command-line Interface

For researchers

```
CUDA_VISIBLE_DEVICES=1,2,3 python "${deepquestpy_dir}/deepquestpy_cli/run_transformer.py" \  
  --model_name_or_path "${pretrained_model_dir}" --arch_name "transformer-word" \  
  --dataset_name "${dataset_name}" --src_lang "en" --tgt_lang "de" \  
  --label_column_name_src "none" --label_column_name_tgt "bad_labels" \  
  --do_train --train_file "${train_file}" \  
  --num_train_epochs 1 --save_steps 10000 --save_total_limit 1 \  
  --learning_rate=5e-6 \  
  --output_dir "${output_dir}" --overwrite_output_dir \  

```



Web Tool for Analysis and Visualisation

For end-users

(1)
Select the languages of
the original (*source*) and
translated (*target*) texts

(3)
Select the model to
use: Transformer or
BiRNN (distilled)

(4)
Choose whether to
use a sentence-level
or word-level model

Enter text

Use the form below the try out the algorithm. When ready, click the "Score" button and the result will appear in a table below the form and be available to download.

Translation Properties

Translation text may be either entered directly using the text boxes below and/or uploaded in a file.

Language

Select the language of the source and target text.

● Source Language English ▼

● Target Language German ▼

Direct Input

Use the text boxes below to submit source-target text pairs to the translation scoring system. If you would like to compare pairs of words or sentences then separate them using line breaks and select the appropriate "Level" setting in the options below.

☒ Source Corpus

☒ Target Corpus

File Upload

Upload translation text in tab-separated values (TSV) format. Do not include column headers.

📎 Corpus File Choose file No file chosen

Options

⚙️ Model

Transformer-based ▼

🔍 Level

Sentence ▼

Score

(2)
Directly type the sentences
to analyse (one per line), or
upload a .tsv file (one
sentence pair per line)

To Summarise...

						
	QuEst	QuEst++	deepQuest	OpenKiwi	TransQuest	deepQuest-py
Large SotA Architectures				✓	✓	✓
Lightweight models	✓	✓	✓	✓		✓
Pre-trained QE Models					✓	✓
Accessible for researchers	✓	✓	✓	✓	✓	✓
Accessible for end-users (e.g. translators)				?		✓

Thanks!

<https://github.com/sheffieldnlp/deepQuest-py>



Fernando Alva-Manchego

 @feralvam

 <https://feralvam.github.io/>



bergam^{ot}

Horizon 2020 Research and Innovation Action
Grant Agreement No. 825303



The
University
Of
Sheffield.