
QUEST: a framework for translation quality estimation

Lucia Specia
Kashif Shah

Department of Computer Science, University of Sheffield
Regent Court, 211 Portobello Street, S4 1DP, UK

l.specia@sheffield.ac.uk
kashif.shah@sheffield.ac.uk

We present QUEST, an open source framework for translation quality estimation. QUEST provides a wide range of feature extractors from source and translation texts and external resources and tools. These go from simple, language-independent features, to advanced, linguistically motivated features. They include features that rely on information from the translation system and features that are oblivious to the way translations were produced. In addition, it provides wrappers for a well-known machine learning toolkit, `scikit-learn`,¹ including techniques for feature selection and model building, as well as parameter optimisation.

We also present a Web interface and functionalities for non-expert users. Using this interface, quality predictions (or internal features of the framework) can be obtained without the installation of the toolkit and the building of prediction models. The interface also provides a ranking method for multiple translations given for the same source text according to their predicted quality.



QuEst - Quality Estimation of Machine Translation

This is a web interface of QuEst that allows users to upload a dataset of source and translations and get quality predictions for a few language pairs for which prediction models have already been pre-built. Features, Predictions and ranking of machine translation are extracted on the fly. This software was developed as part of the QUEST and QTLaunchPad projects.

The screenshot shows two panels from the QuEst web interface. The left panel, titled "Get Translation, Features and Prediction", provides instructions for file upload. It states: "Select a file to upload: Please make sure that file has an extension .txt and contain a sentence per line and is in following format with ||| as delimiter: id ||| source sentence to be translated e.g. 1 ||| this is first sentence 2 ||| this is second sentence If input sentences are already translated, please add additional field as follows: id ||| source sentence to be translated ||| translated sentence e.g. 1 ||| this is first sentence ||| THIS IS FIRST SENTENCE TRANSLATION 2 ||| this is second sentence ||| THIS IS SECOND SENTENCE TRANSLATION". Below the text are "Browse..." and "Upload File" buttons. The right panel, titled "Get Ranking of multiple targets given same source:", shows a "Format (tab separated):" section with a table: "1 source-sentence target1 target2 target3". Below the table are "Browse..." and "Upload File and Rank" buttons, and a "Choose the language pair:" dropdown menu.

Acknowledgements

This work was supported by funding from the from European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 296347 (QTLaunch-Pad).

¹<http://scikit-learn.org/>