

On the Usability of Transformers-based models for a French Question-Answering task - abstract

Oralie Cattan^{1,2,*}, Christophe Servan^{1,2,†} and Sophie Rosset^{2,†}

¹Paris-Saclay University, CNRS, LISN

²QWANT

Abstract

Transformers have sparked a paradigmatic shift in question-answering training practices by simplifying its architectures. As models became larger and better important usability shortcomings appeared. This includes their computational costs and degraded performance with limited training data (e.g., domain-specific or low-resourced language tasks). Considering this resource trade-off, we (i) explore training strategies such as data augmentation, hyperparameter optimization, cross-lingual transfers and cross-dataset mixing, (ii) perform an in-depth analysis to understand the contribution of each on model performance maintenance and (iii) provide a question-answering corpus and a compressed pre-trained model for French¹. Our experimental results attest to the merit of a flexible paradigm for a low-resource scenario.

Keywords

question answering, transformer architectures, pre-trained models and scalability, language resources

¹Both are available at: <https://huggingface.co/qwant>

CIRCLE (Joint Conference of the Information Retrieval Communities in Europe) 2022, July 04–07, 2022, Samatan, France

*Corresponding author.

†These authors contributed equally.

✉ oralie.cattan@lisn.fr (O. Cattan); christophe.servan@lisn.fr (C. Servan); sophie.rosset@lisn.fr (S. Rosset)

🆔 0000-0003-2805-5620 (O. Cattan); 0000-0003-2306-7075 (C. Servan); 0000-0002-6865-4989 (S. Rosset)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)