CLAfICLe: Cross Lingual Adaptation for In-Context Learning

Giulio Starace
University of Amsterdam / Amsterdam, The Netherlands
giulio.starace@gmail.com

Abstract

This document is a supplement to the general instructions for *ACL authors. It contains instructions for using the LaTeX style files for ACL conferences. The document itself conforms to its own specifications, and is therefore an example of what your manuscript should look like. These instructions should be used both for papers submitted for review and for final versions of accepted papers.

1 Introduction

2 Related Work

3 Method

4 Results and Discussion

Table 1: Average performance across the datasets from our multi-task benchmark for the models considered in this work. We report average difference in performance for each proposed alternative to Sandwich. Negative values indicate underperformance compared to Sandwich.

	en	fr	de
MetaICL	0.327	-	-
Sandwich	-	0.317	0.322
Difference in Performance w.r.t. Sandwich			
MetaICL-W	-	-0.020	-0.026
GPT2-W+MetaICLA	-	-0.041	-0.042
GPT2-W+MetaICLVA	-	-0.036	-0.045

5 Conclusion

References

Sewon Min, Mike Lewis, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2022. MetaICL: Learning to Learn In Context. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2791–2809, Seattle, United States. Association for Computational Linguistics.

Benjamin Minixhofer, Fabian Paischer, and Navid Rekabsaz. 2022. WECHSEL: Effective initialization of subword embeddings for cross-lingual transfer of monolingual language models. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3992–4006, Seattle, United States. Association for Computational Linguistics.

6 Appendices

Use `\appendix` before any appendix section to switch the section numbering over to letters. See Appendix 7 for an example.

7 Example Appendix

This is an appendix.

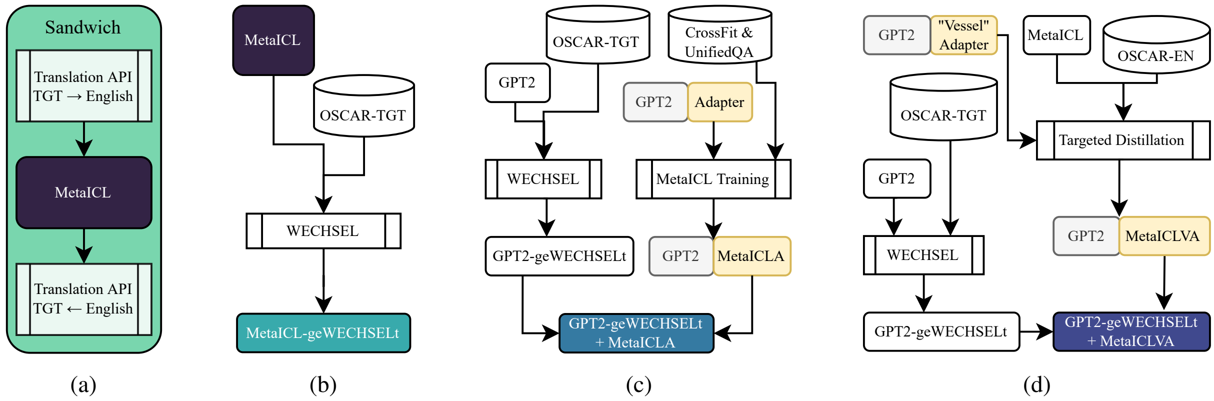

Figure 1: Overview of each of the models evaluated in one of the two TGT languages (French or German). The baseline Sandwich model (a) sandwiches MetaICL (Min et al., 2022) (which we separately evaluate only in English) between two complementary translation API calls. MetaICL-geWECHSELt (b) is the result of applying WECHSEL (Minixhofer et al., 2022) to MetaICL. GPT2-geWECHSELt+MetaICLA combines MetaICLA, an adapter trained on the MetaICL dataset and objective, with a TGT-language GPT2 base obtained via WECHSEL. GPT2-geWECHSELt+MetaICLVA does the same, except MetaICLVA is trained via targeted distillation with supervision provided by MetaICL. For more details, refer to section 3.

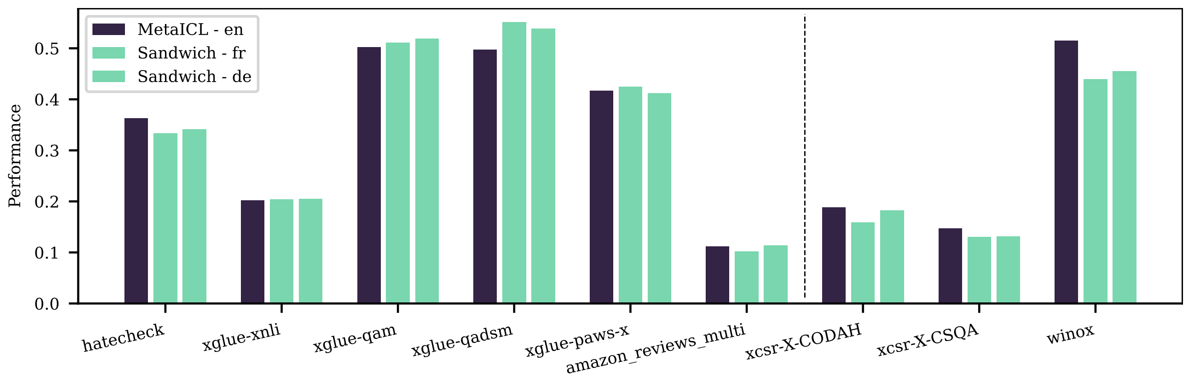

Figure 2: Performance on a particular language dimension of our multi-task benchmark of our two baseline models, MetaICL and Sandwich. The dashed line separates whether a given task uses accuracy (left) or F1-score (right) as the performance metric.

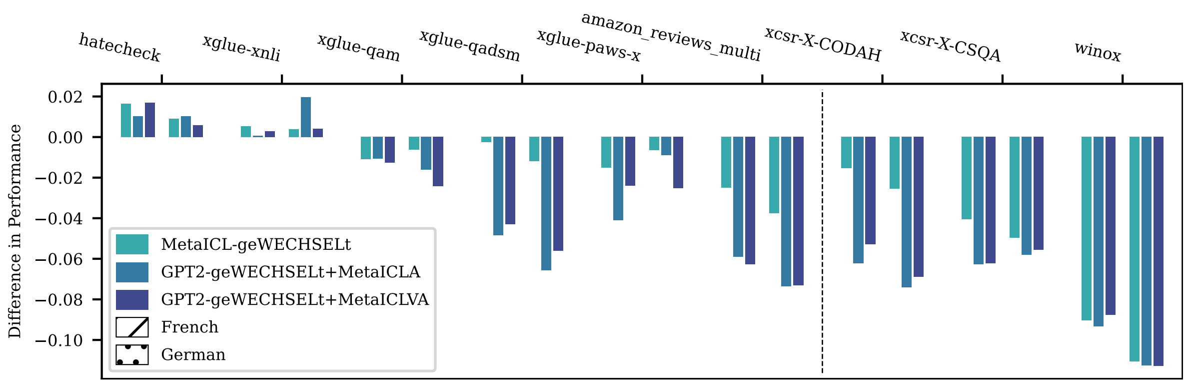

Figure 3: Performance gap on our multi-task benchmark between each of the language-adapted models and the "Sandwich" baseline. Positive values indicate that the adapted models are outperforming the baseline, while negative values indicate the reverse. The dashed line separates whether a given task uses accuracy (left) or F1-score (right) as the performance metric.