Cross-language continual learning for conversational systems

Sahar Ghannay, Laure Soulier, Christophe Servan, Sophie Rosset
LISN Laboratory - Saclay

Subject

Understanding natural language in the context of conversational systems is a critical step to ensure their effectiveness. Different work in natural language processing community are devoted to solve natural language understanding (NLU), tackling various tasks such as information extraction, paraphrase or summarization, or also slot-filling. Mots of these works have been possible thanks to the introduction of large language models which have demonstrated large capabilities over different tasks. However, one drawback of these solutions is that they are often addressed for a specific language or a small fraction of languages. If we desire to deploy virtual assistants over all the world, it is therefore important to design models able to address a large number of languages.

In this internship, we assume that the deployment of virtual assistants can be done step by step over different countries in the world and, thus, that virtual assistants will face different languages at different timestamp. This assumption imply that, when designing/training a model for a given task, languages can be incrementally added in the training procedure. This setting refers to as two main research fields:

- Cross-lingual transfer [Coria et al., 2022], which aims at exploiting the knowledge of languages previously exploited in a pre-training process to train the model on another language. In such setting, the knowledge from previous language will serve as initialization of the language model of another language, enabling to reduce the training time.

- Continual learning [Kirkpatrick et al., 2016, Ke et al., 2020], which aims at designing models trained on a stream of tasks by learning knowledge from new tasks without forgetting what has been learned on previous ones. Some work have been proposed in NLP [Lee, 2017, Garcia et al., 2021].

In our case, we propose a continual learning setting in which the task is fixed, but the stream is based on different languages. The model therefore learn the knowledge of language pecu-

liarities. Therefore, to satisfy the initial condition of virtual assistants to address different languages, we therefore need to ensure that our task-based model does not forget previous languages while training on new ones.

Two preliminary works have been done: 1) [Coria et al., 2022][1], investigating BERT's cross-lingual transfer capabilities in two continual sequence labeling tasks. 2) [Gerald and Soulier, 2022] designing continual learning streams for information retrieval.

In practice, we will focus on the Massively Multilingual NLU 2022 data [FitzGerald et al., 2022], which includes slot-filling and NER tasks for 51 languages in parallel. The objective of the internship will be to 1) build a stream of languages for a given task, 2) run baseline models in the stream, and 3) design a continual learning model for cross-lingual transfer.

Information

Supervisors: Sahar Ghannay, Laure Soulier, Christophe Servan, Sophie Rosset
Contact: sahar.ghannay@lisn.fr, laure.soulier@isir.upmc.fr, christophe.servan@lisn.fr, sophie.rosset@lisn.fr
Localization: Université Paris Saclay (Laboratoire LISN), France
Duration: 6 months, between February and August 2023.
Stipend: around 591.91 euros / month[2]

Expected profile: Master or engineering degree in Computer Science or Applied Mathematics related to machine learning/natural language processing. The candidate should have a strong scientific background with good technical skills in programming, and be fluent in reading and writing English. Autonomy and curiosity are also adequate soft-skills to work on this internship.

How to apply? Send a CV, a motivation letter and Master records to `sahar.ghannay@lisn.fr`, `laure.soulier@isir.upmc.fr`, `christophe.servan@lisn.fr`, `sophie.rosset@lisn.fr`. Recommendation letters would be appreciated.
Interviews will conducted as they arise and the position will be filled as soon as possible – the latest application date is set to 15th January.

References

[Coria et al., 2022] Coria, J. M., Veron, M., Ghannay, S., Bernard, G., Bredin, H., Galibert, O., and Rosset, S. (2022). Analyzing BERT cross-lingual transfer capabilities in continual sequence labeling. In *Proceedings of the First Workshop on Performance and Interpretabil-*

---

[1]https://github.com/juanmc2005/ContinualNLU
[2]Standard internship stipend in France – Computed on Government Website: `https://www.service-public.fr/simulateur/calcul/gratification-stagiaire`, new law to be published that should make it higher in 2023 but actual figures are yet unknown.

ity Evaluations of Multimodal, Multipurpose, Massive-Scale Models, pages 15–25, Virtual. International Conference on Computational Linguistics.

[FitzGerald et al., 2022] FitzGerald, J., Hench, C., Peris, C., Mackie, S., Rottmann, K., Sanchez, A., Nash, A., Urbach, L., Kakarala, V., Singh, R., Ranganath, S., Crist, L., Britan, M., Leeuwis, W., Tur, G., and Natarajan, P. (2022). Massive: A 1m-example multilingual natural language understanding dataset with 51 typologically-diverse languages.

[Garcia et al., 2021] Garcia, X., Constant, N., Parikh, A. P., and Firat, O. (2021). Towards continual learning for multilingual machine translation via vocabulary substitution. In *NAACL-HLT*, pages 1184–1192.

[Gerald and Soulier, 2022] Gerald, T. and Soulier, L. (2022). Continual learning of long topic sequences in neural information retrieval. In Hagen, M., Verberne, S., Macdonald, C., Seifert, C., Balog, K., Nørvåg, K., and Setty, V., editors, *Advances in Information Retrieval - 44th European Conference on IR Research, ECIR 2022, Stavanger, Norway, April 10-14, 2022, Proceedings, Part I*, volume 13185 of *Lecture Notes in Computer Science*, pages 244–259. Springer.

[Ke et al., 2020] Ke, Z., Liu, B., and Huang, X. (2020). Continual learning of a mixed sequence of similar and dissimilar tasks. In *NeurIPS*.

[Kirkpatrick et al., 2016] Kirkpatrick, J., Pascanu, R., Rabinowitz, N. C., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., and Hadsell, R. (2016). Overcoming catastrophic forgetting in neural networks. *CoRR*, abs/1612.00796.

[Lee, 2017] Lee, S. (2017). Toward continual learning for conversational agents. *CoRR*, abs/1712.09943.