

Évaluer les modèles de langue pré-entraînés avec des propriétés de hiérarchie

Jesús Lovón-Melgarejo¹ Jose G. Moreno¹ Romaric Besançon² Olivier Ferret²
Lynda Tamine¹

(1) Université Paul Sabatier, IRIT, Toulouse, France

(2) Université Paris-Saclay, CEA, List, Palaiseau, France

{jesus.lovon, jose.moreno, tamine}@irit.fr,
{romaric.besancon, olivier.ferret}@cea.fr

RÉSUMÉ

Étant donné que les modèles de langue pré-entraînés (PLM) constituent la pierre angulaire des modèles de recherche d'informations les plus récents, la façon dont ils encodent la connaissance sémantique est particulièrement importante. Cependant, on s'est peu intéressé à la capacité des PLM à capturer la connaissance sémantique hiérarchique. Traditionnellement, l'évaluation de ces connaissances codées dans les PLM s'appuie sur leurs performances lors d'évaluations dépendantes de la tâche, basées sur des tâches proxy telles que la détection d'hyperonymes. Malheureusement, cette approche ignore potentiellement d'autres relations taxonomiques implicites et complexes. Dans ce travail, nous proposons une méthode d'évaluation indépendante de la tâche, capable d'évaluer dans quelle mesure les PLM peuvent capturer des relations taxonomiques complexes, telles que les ancêtres et les frères et sœurs. Cette évaluation, basée sur des propriétés intrinsèques capturant ces relations, montre que les connaissances lexico-sémantiques codées implicitement dans les PLM ne capturent pas toujours les relations hiérarchiques. Nous démontrons en outre que les propriétés proposées peuvent être injectées dans les PLM pour améliorer leur compréhension de la hiérarchie. Grâce à des évaluations portant sur la reconstruction de taxonomies, la découverte d'hyperonymes et la compréhension de lecture, nous montrons que la connaissance de la hiérarchie est modérément transférable entre les tâches, mais pas de manière systématique. Ceci est le résumé de l'article "Probing Pretrained Language Models with Hierarchy Properties" publié à ECIR 2024 (Lovón-Melgarejo *et al.*, 2024).

ABSTRACT

Probing Pretrained Language Models with Hierarchy Properties

Since Pretrained Language Models (PLMs) are the cornerstone of the most recent Information Retrieval models, the way they encode semantic knowledge is particularly important. However, little attention has been given to studying the PLMs' capability to capture hierarchical semantic knowledge. Traditionally, evaluating such knowledge encoded in PLMs relies on their performance on task-dependent evaluations based on proxy tasks, such as hypernymy detection. Unfortunately, this approach potentially ignores other implicit and complex taxonomic relations. In this work, we propose a task-agnostic evaluation method able to evaluate to what extent PLMs can capture complex taxonomy relations, such as ancestors and siblings. This evaluation, based on intrinsic properties capturing these relations, shows that the lexico-semantic knowledge implicitly encoded in PLMs does not always capture hierarchical relations. We further demonstrate that the proposed properties can be injected into PLMs to improve their understanding of hierarchy. Through evaluations on taxonomy reconstruction, hypernym discovery and reading comprehension tasks, we show that knowledge about

hierarchy is moderately but not systematically transferable across tasks. This is the summary of the published paper "Probing Pretrained Language Models with Hierarchy Properties" at ECIR 2024 (Lovón-Melgarejo *et al.*, 2024).

MOTS-CLÉS : modèles neuronaux de langue, relations taxonomiques, évaluation .

KEYWORDS: neural language models, taxonomic relations, evaluation .

Références

LOVÓN-MELGAREJO J., MORENO J. G., BESANÇON R., FERRET O. & TAMINE L. (2024). Probing pretrained language models with hierarchy properties. In *Advances in Information Retrieval : 46th European Conference on IR Research, ECIR 2024, Glasgow, Scotland, March 24–March 28, 2024* : Springer International Publishing.