
Master thesis : Sparse hypernetworks for multitasking

Auteur : Cubélier, François

Promoteur(s) : Geurts, Pierre

Faculté : Faculté des Sciences appliquées

Diplôme : Master en ingénieur civil en informatique, à finalité spécialisée en "intelligent systems"

Année académique : 2021-2022

URI/URL : <http://hdl.handle.net/2268.2/14574>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Sparse hypernetworks for multitasking

Summary

Author: **Francois Cubelier**
Advisor: **Prof. Pierre Geurts**

Master's thesis completed in order to obtain the
Degree Of Science (MSc) in Computer Science and Engineering

Machine learning researchers have always been interested in creating less narrow artificial intelligence. Meta-models, i.e. models capable of producing other models, could potentially be a key ingredient for building new highly multitasking capable models. Hypernetworks, which are neural networks that produce the parameters of other neural networks, can be used as meta-models. However, due to the large number of parameters in neural networks nowadays, it is not trivial to build hypernetworks with the large output size required to produce all the parameters of another neural network. Current solutions, like chunked hypernetworks, which split the target parameter space into parts and reuse the same model to produce each part, achieve good results in practice and are scalable independently of the maximal size of the layers in the target model. However, they seem unsatisfactory because they arbitrarily split the target model parameters into chunks. In this work, we propose a new scalable architecture for building hypernetworks, which consists in a sparse MLP with hidden layers of exponentially growing size. After testing different variations of this architecture, we compare it with chunked hypernetworks on multitasking computer vision benchmarks. We show that they can match the performance of chunked hypernetworks, even though they were slightly behind on more complex problems. We also show that linear sparse hypernetworks outperformed their non-linear version and chunked hypernetworks for inferring new models for new tasks with a pretrained task-conditioned hypernetwork. This may indicate that linear sparse hypernetworks have better generalization properties than more complex hypernetworks. In addition to proposing this sparse architecture and as a preamble of this work, we also review the literature on hypernetworks and propose a typology of hypernetworks. Even though the results obtained are promising, there are still many ways to improve sparse hypernetworks and, more generally, hypernetworks that can be explored in future research.



University of Liège - School of Engineering
Belgium
Academic year 2021-2022