

Apprentissage multi-tâches, apprentissage par transfert : adaptabilité des réseaux de neurones

Romain Orhand / rorhand@unistra.fr
Anne Jeannin-Girardon / anne.jeannin@unistra.fr

Contexte

Si les modèles d'apprentissage sont de plus en plus précis et efficaces, surpassant parfois les performances humaines (transcription de la parole, analyse d'images, etc.), il est nécessaire de disposer de données d'entraînement de bonne qualité et en quantité importante. Selon les domaines ça n'est pas toujours le cas, comme par exemple dans le domaine de l'imagerie médicale où les données peuvent être soit moins nombreuses que dans d'autres tâches de reconnaissance d'images, soit moins (bien) étiquetées car cela nécessite l'intervention d'experts humain. Sans aller dans des domaines aussi spécifiques que l'imagerie médicale, on peut aussi prendre pour exemple l'idée d'appliquer un algorithme d'analyse de sentiment entraîné sur des avis de films à des avis de livres, qui sont deux tâches en relations, bien qu'elles soient différentes : pouvoir ré-appliquer un modèle déjà entraîné à une tâche sur une tâche similaire permettrait 1) de pallier un éventuel manque de données liées à la nouvelle tâche et 2) de réduire la durée d'apprentissage pour la nouvelle tâche puisqu'on ne démarre pas avec des paramètres aléatoires.

L'apprentissage par transfert (utiliser un modèle entraîné sur une tâche donnée à une tâche différente mais corrélée à la première) et l'apprentissage multi-tâches (apprentissage de plusieurs tâches similaires de façon simultanée) sont des pistes prometteuses pour améliorer les performances des algorithmes d'apprentissage. Les deux enjeux principaux de ces techniques sont, d'une part, de pouvoir entraîner un modèle avec moins de données et, d'autre part, de réduire le temps d'entraînement des modèles. Ainsi, en utilisant de telles techniques d'apprentissage, il serait possible d'obtenir des modèles plus adaptatifs puisqu'ils seraient mieux capable de généraliser (là où les modèles actuels sont très spécifiques à la tâche sur laquelle ils ont été entraînés et donc difficilement réutilisables pour d'autres tâches, même similaires). Cette adaptabilité peut être issue, à titre d'exemple, de l'intégration de modèles pré-entraînés [1][2], de l'entraînement d'un modèle sur des tâches distinctes simultanément [3][4], de l'extraction et exploitation de caractéristiques propres aux tâches [5][6].

Travail demandé

L'objectif de ce travail est en premier lieu de réaliser une étude des différentes pistes et techniques pour faire de l'apprentissage par transfert/multi-tâches selon les contextes (domaines, tâches / transfert de caractéristiques, transfert de paramètres). En second lieu, il faudra aussi se poser la question de comment il serait possible d'évaluer les similarités entre des tâches et des domaines au moyen d'outils statistiques (matrices de covariance) ou d'outils représentatifs tels que les diagrammes de Hinton. Il pourra être intéressant dans le même temps d'explorer de nouvelles possibilités afin de mettre en avant ces similitudes en recherchant de nouvelles représentations des modèles utilisés, qui seraient alors en mesure de mettre en évidence de nouvelles informations.

Références

- [1] Dahl, D. et al., "Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 1, pp. 30-42, Jan. 2012. doi: 10.1109/TASL.2011.2134090
- [2] Erhan et al., "Why Does Unsupervised Pre-training Help Deep Learning", JMLR, 2010, 10.1145/1756006.1756025
- [3] Marquand et al., "Bayesian multi-task learning for decoding multi-subject neuroimaging data", in Neuroimage, 2014, doi: 10.1016/j.neuroimage.2014.02.008
- [4] Luan et al., "Multi-Task Learning for Speaker-Role Adaptation in Neural Conversation Models", arXiv:1710.07388 [cs.CL], 2017
- [5] Bengio, "Deep Learning of Representations for Unsupervised and Transfer Learning", in JMLR 2012, <http://proceedings.mlr.press/v27/bengio12a/bengio12a.pdf>
- [6] Yosinski et al., "How transferable are features in deep neural networks?" in NIPS 2014, <https://arxiv.org/abs/1411.1792>
- [7] Brownlee, "A Gentle Introduction to Transfer Learning for Deep Learning", Dec. 2017, <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
- [8] Ruder, "Transfer Learning - Machine Learning's Next Frontier", March 2017, <http://ruder.io/transfer-learning/>