

Recrutement post-doctorat (1 ans – Cnam Paris)

Explicabilité des modèles transformers pour la vision et la recommandation.

1 Contexte

Les réseaux de neurones basés sur des blocs transformers ont récemment connus de nombreux succès en traitement naturel du langage [1] ou en vision par ordinateur [2, 4]. Ces architectures ont la particularité de s'appuyer sur un mécanisme d'auto-attention mettant en relation l'ensemble des éléments en entrée du réseau et peuvent ainsi extraire des dépendances sur plusieurs niveaux d'échelle. Cette auto-attention présente la possibilité d'être ré-interprétée comme un graphe d'interactions [3] et ouvre des pistes prometteuses dans la construction de modèles explicables.

2 Objectifs

Le travail effectué par la personne recrutée sera décliné à travers deux axes : i) un travail sur l'architecture et l'entraînement de modèles transformers pour l'apprentissage de prototypes ou de représentations génériques, ainsi que ii) l'exploitation et la modélisation des relations entre ces prototypes et les entrées pour caractériser ces dernières.

Les travaux du candidat devront s'articuler autour des deux applications : la segmentation d'image et le système de recommandation. Sur ces deux points le candidat viendra en renfort pour épauler des doctorants associés à des projets démarrés au cours de l'année précédente.

3 Profil

Nous recherchons un-e candidat-e ayant soutenu une thèse de doctorat en apprentissage automatique, mathématiques appliquées, traitement de l'image et/ou du signal. Le ou la candidat-e idéal-e a une appétence pour la recherche scientifique et des bases solides en apprentissage automatique.

4 Organisation

Cette offre de postdoc est d'une durée de 12 mois avec une date de début flexible à l'automne 2022. Le postdoc sera effectué au centre de recherche et d'études en informatique et en communications (CEDRIC) du Conservatoire national des arts et métiers (Cnam) à Paris, 3^e arrondissement.

Les travaux du postdoctorant-e seront effectués en collaboration avec Clément Rambour, Raphaël Fournier et Nicolas Thome de l'équipe **Données complexes, apprentissage et représentations**.

5 Candidater

Envoyer une candidature (CV + brève explication de votre motivation) par email à clement.rambour@cnam.fr.

Références

- [1] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words : Transformers for image recognition at scale, 2020.
- [2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need, 2017.
- [3] S. Yun, M. Jeong, R. Kim, J. Kang, and H. J. Kim. Graph transformer networks. *Advances in neural information processing systems*, 32, 2019.
- [4] P. Zhang, X. Dai, J. Yang, B. Xiao, L. Yuan, L. Zhang, and J. Gao. Multi-scale vision longformer : A new vision transformer for high-resolution image encoding, 2021.