Recrutement d'un allocataire post-doctoral

AVIS D APPEL A CANDIDATURES

Post-doc POST OFFER

1. Descriptif du poste proposé / Offer description

Poste de catégorie A / Position level: A

Etablissement d'emploi / Employment establishment	Université du Mans Avenue Olivier Messiaen 72085 LE MANS CEDEX 11
Laboratoire d'exercice des fonctions / Research laboratory	LIUM - Laboratoire d'informatique de l'Université du Mans - EA403
Métier auquel se rattache l'emploi proposé / Job	Chercheur/Chercheuse (détail consultable à ce lien: https://www.fonction-publique.gouv.fr/chercheusechercheur) / searcher
Intitulé du poste/ Job title:	Transformers interprétables / Interpretable transformers

Description du projet ou de l'opération de recherche dans lesquels s'inscrivent les activités de recherche confiées / Description of the research subject

En prenant comme première brique les représentations interprétables apprises avec SINr, l'approche de plongements interprétables comme première brique, le ou la candidat·e recruté·e aura la charge des architectures neuronales de classification interprétables de bout-en-bout. L'objectif est de rester dans un espace interprétable tout au long de la classification. Ainsi, des mécanismes profonds pourront être mis en œuvre en se basant sur la structure hiérarchique des plongements produits par SINr, et en s'inspirant par exemple des travaux de Victoria Bourgeais [BZBHH21]. Des mécanismes d'attention de type dot product tels que dans Bahdanau [BCB14], utilisant un vecteur d'attention dédié à la tâche, qui, s'il est dans le même espace que l'entrée sera également interprétable. Mais d'autres approches sont également envisageables pour exploiter l'interprétabilité au sein de modèles plus complexes tels que les transformers. Clark et al. [CKLM19] ont mis en évidence les rôles joués par les têtes d'attention, et notamment leur spécialisation. Geva et al. [GSBL20] ont travaillé sur les modules feed-fordward du transformer pour déterminer leur importance. Enfin, Mickus et al. [MPC22] dissèquent le transformer pour mesurer la contribution de chacun de ses modules (attention, biais, feed-forward, embedding de départ) dans les représentations de sortie mais également dans la prédiction du mot masqué. Ainsi, l'état de l'art a progressé sur l'explicabilité des transformers et de leurs mécanismes, permettant ainsi d'envisager des architectures réduites et interprétables s'en inspirant. Pour l'évaluation de ces architectures, nous envisagerons des tâches de classification telles que la reconnaissance d'entités nommées, l'analyse de polarité ou la détection de contenus haineux. Mais il s'agira également de développer un cadre d'évaluation de l'interprétabilité bout-en-bout.

Based on the representations learned with SINr, with the interpretable plunging approach as the first building block, the candidate recruited will be in charge of end-to-end interpretable classification neural architectures. The aim is to remain in an interpretable space throughout the classification. In this way, deep mechanisms can be implemented based on the hierarchical structure of the dives produced by SINr, and inspired for example by the work of Victoria Bourgeais [BZBHH21]. Attention mechanisms of the dot product type as in Bahdanau [BCB14], using an attention vector dedicated to the task, which, if it is in the same space as the input, will also be interpretable. But other approaches are also possible for exploiting interpretability within more complex models such as transformers. Clark et al [CKLM19] have highlighted the roles played by attention heads, and in particular their specialisation. Geva et al [GSBL20] worked on the feed-fordward modules of the transformer to determine their importance. Finally, Mickus et al [MPC22] dissected the transformer to measure the contribution of each of its modules (attention, bias, feed-forward, initial embedding) in the output representations and also in the prediction of the hidden word. In this way, the state of the art has made progress on the explicability of transformers and their mechanisms, allowing us to envisage reduced and interpretable architectures inspired by them. To evaluate these architectures, we will consider classification tasks such as named entity recognition, polarity analysis or hate content detection. But it will also involve developing an end-to-end interpretability evaluation framework.

Calendrier prévisionnel du projet ou de l'opération de recherche / Planning research project

Etat de l'art, prise en main des transformers, développement des modèles interprétables, évaluation

Literature review, coding transformers, upgrading the code to make them interpretable, evaluation

Activités de recherche confiées, tâches à accomplir et résultats attendus / assigned activities and expected results

Etat de l'art, prise en main des transformers, développement des modèles interprétables, évaluation

Literature review, coding transformers, upgrading the code to make them interpretable, evaluation

Qualification requises pour l'exercice des fonctions / Required education level	- être titulaire du doctorat / PhD or equivalent - autres qualifications: -

Compétences attendue / Candidate profile

Python, pytorch, github, transformers

Python, pytorch, github, transformers

Mots clefs / Key Words

Computer science Informatique

Conditions d'exercice des missions / working conditions	Sur site au Mans At Lium, in Le Mans

Eventuelles sujetions particulières attachées au poste	-
Localisation géographique du poste / Work location	Le Mans
Prise de fonction souhaitée le / offer starting date:	15-juil-24
Durée prévue des missions confiées (durée initiale du contrat proposé) / contract length	1 an
Fin de fonction prévue le / offer ending date:	14-juil-25
Rémunération brute mensuelle proposée / Salary	2 800 €
Quotité du poste / Work quota	100% / Full time

L'emploi est-il financé dans le cadre d'un Programme cadre de recherche de l'UE ? / Is the job funded through a EU Research FrameworkProgramme?

Non financé par un programme de l'UE
Not funded by an EU programme

Numéro de référence / Reference Number : -

Si le candidat a obtenu son doctorat depuis moins de trois ans, il se verra proposé un contrat de postdoctorant de droit public prévu par les dispositions du Décret n°2021-1450. Le cas échéant, il aura un contrat équivalent. / If the candidate obtained their doctorate less than three years ago, he will have a public law postdoctoral contract provided for by the provisions of Decree No. 2021-1450.

2. Description du processus de recrutement / recruitment process

Documents à fournir à l'appui de la candidature / Documents to provide

- Curriculum vitae détaillé / detailed curriculum vitae
- Lettre de motivation / cover letter
- diplôme de doctorat / doctoral degree
- autres pièces: -

Contact pour plus d'informations sur le poste proposé et candidatures à adresser à / Contact for more informations on the position and applications to be sent to:

Le Mans Université
LIUM - Laboratoire d'informatique de l'Université du Mans - EA403
M. Nicolas Dugué
mail: nicolas.dugue@univ-lemans.fr
Tel: 0616422489

Contact administratif/ Administrative contact:

Le Mans Université
Direction des Ressources Humaines-
Pôle gestion des personnels enseignants-chercheurs
Mme Ombline Cador / Mme Jennifer Leboucher
drh-post-doc@univ-lemans.fr
Tel: 02 43 83 39 37 / 02 43 83 26 70

Date et heure limites de dépôt des candidatures / Deadline for submitting applications: **7 juillet 2024 à 16 h**

Le présent recrutement fera l'objet d'un entretien de recrutement / This recruitment will be the subject of a recruitment interview