

Contrat post-doctoral H/F : IA explicable, IA de confiance, IA pour l'éducation

Laboratoire de recherche : LIUM (Laval)

Encadrement : Rémi Venant (LIUM), Marie Lefevre et Nathalie Guin (LIRIS)

Contacts : Remi.Venant@univ-lemans.fr, marie.lefevre@univ-lyon1.fr,
nathalie.guin@univ-lyon1.fr

Durée du contrat : 1 an

Date d'embauche : dès que possible

Contexte

Ce poste s'inscrit dans le projet **ANR TEA-TIME (Teacher's Technology Integration Models in Education¹)**, dont l'objectif est de fournir des modèles et des moyens pour l'utilisation d'activités fondées sur l'approche par compétences (APC) utilisant des régulations d'Intelligence Artificielle pour favoriser l'auto-régulation.

Dans un projet ANR précédent (COMPER²), nous avons produit des outils permettant de mettre en œuvre une approche par compétences pour accompagner l'apprentissage de manière personnalisée [1]. L'objectif du projet TEA-TIME est de compléter les possibilités fonctionnelles du cadre COMPER en permettant d'expérimenter la mise en œuvre d'une large variété de dispositifs de formation fondés sur l'APC. Pour cela, le projet vise à contribuer au soutien à l'intégration du numérique dans les pratiques des enseignants par une analyse de la recevabilité (utilité potentielle, acceptabilité, appropriabilité) des solutions proposées.

Le projet s'appuie sur un partenariat entre deux laboratoires SHS, trois laboratoires d'informatique et une entreprise éditeur d'ENT. Ils collaborent à la conception des services qui complètent COMPER (planification, supervision, explication et diagnostic) et à l'analyse de leurs usages. Les contributions en informatique portent sur l'explication des méthodes d'IA qui sont sous-jacentes aux services ciblés.

Missions

L'objectif du projet ANR COMPER était de concevoir un cadre permettant aux enseignants de concevoir et orchestrer des activités permettant de mettre en œuvre l'APC, en adaptant celles-ci aux objectifs pédagogiques des enseignants et à leurs apprenants. Pour cela, différents moyens et modèles ont été proposés : un méta-modèle de référentiel de compétences et deux systèmes d'IA exploitant ce méta-modèle. Le premier système d'IA est un moteur d'analyse des traces d'activités pédagogiques permettant de construire un profil de compétences (i.e.: la maîtrise qu'a l'étudiant des différentes compétences) et le second un moteur de recommandation de ressources pédagogiques [2].

Ces systèmes d'IA s'appuient sur des approches d'IA symbolique et peuvent être amenés à prendre des décisions incorrectes, de part un raisonnement faussé ou un paramétrage

¹ <https://anr.fr/Projet-ANR-23-CE38-0012>

² <https://comper.fr/>

utilisateur inadapté. Si la notion de précision est très souvent utilisée pour estimer à quel point un modèle fournit des solutions correctes, celle-ci n'est pas un indicateur satisfaisant pour des systèmes critiques où l'humain doit être en mesure de comprendre l'estimation du modèle (avant de prendre une décision ou pour le surveiller).

Le traitement de cette problématique s'appuie principalement sur l'explicabilité, qui vise à obtenir des informations pertinentes sur la prise de décision d'un modèle au niveau global comme au niveau local [3-4], notamment par la fabrication d'exemples contrefactuels, de tests aléatoires ou par des approches formelles [5]. Dans le domaine de l'éducation, cette problématique se retrouve également dans l'étude des biais algorithmiques sur l'apprentissage [6]. Ces différentes recherches font consensus sur l'inexistence de méthodes de vérification universelle, la dépendance de l'approche à la spécificité du contexte, ainsi que le manque de travail empirique sur les conséquences des biais algorithmiques sur l'apprentissage.

Il est donc crucial de mettre en œuvre des méthodes d'explications et de diagnostic permettant aux enseignants ou ingénieurs pédagogiques de vérifier les décisions prises par les modèles d'IA sous-jacents et aux apprenants de comprendre sur quoi repose leur profil de compétences pour soutenir les processus d'auto-régulation de l'apprentissage.

Les premiers résultats du projet TEA-TIME ont permis de proposer des mécanismes d'explications, d'une part à destination des enseignants concernant les décisions prises par le moteur de recommandation, et d'autre part à destination des apprenants concernant le calcul des taux de maîtrise du profil. Ces explications sont générées à la volée en s'appuyant sur les données à l'origine du calcul (les traces des exercices réussis ou échoués), sur la structure du référentiel de compétences et l'ontologie le modélisant faisant intervenir des relations sémantiques entre les différents niveaux de compétences, savoirs et savoir-faire, mais également sur les traces des moteurs d'IA permettant de connaître les raisons des décisions prises.

La mission consiste à **compléter les deux moteurs d'explication** pour permettre une explication complète des deux outils, et de proposer des **mécanismes d'adaptation de la présentation de ces explications** en fonction de la personne qui les reçoit (enseignant / apprenant). Cela devrait favoriser l'appropriation de ces deux outils par les enseignants, pour leur permettre de les paramétrer selon leurs besoins, et favoriser leur acceptation par les apprenants.

Activités

- comprendre le processus de calcul des profils de compétences des apprenants
- comprendre le processus de personnalisation mis en œuvre
- compléter et généraliser les processus d'explication de ces deux systèmes d'IA pour les enseignants, en lien avec les paramétrages qu'ils peuvent réaliser
- proposer un mécanisme d'adaptation pour présenter les explications aux apprenants
- analyser l'usage de ces outils par les enseignants et les apprenants

- participer à la rédaction d'articles scientifiques et à la présentation des résultats de la recherche

Compétences attendues

Nous recherchons une personne :

- diplômé.e d'un doctorat en informatique dans l'un des domaines suivants : EIAH, Ingénierie des Connaissances, IHM, IA
- qui s'intéresse aux technologies du numérique au service de l'éducation ;
- qui possède
 - de solides compétences dans le développement web ;
 - une expérience pratique dans le domaine du développement logiciel ;
 - des connaissances dans le domaine de l'explicabilité de l'IA et/ou de l'appropriation des outils numériques par les utilisateurs ;
- qui est habituée à collaborer avec d'autres chercheurs ;
- qui a un bon niveau d'anglais (lecture, écriture et expression orale).

Contexte de travail

La personne recrutée rejoindra l'équipe IEAIH du laboratoire LIUM, située sur le campus de d'IUT Laval (CERIUM 2) et travaillera sous la direction de Rémi Venant (LIUM), en collaboration avec Marie Lefevre et Nathalie Guin (LIRIS).

[1] L. Pierrot, C. Michel, J. Broisin, N. Guin, M. Lefevre, et R. Venant, « Assessing COMPER Environment to Support Self-Regulation During Autonomous Work », in *Open and Inclusive Educational Practice in the Digital World*, 2023, p. 193-212.

[2] L. Sablayrolles, M. Lefevre, N. Guin, et J. Broisin, « Design and Evaluation of a Competency-based Recommendation Process », in *Intelligent Tutoring Systems*, Bucharest, Romania, 2022.

[3] O. Biran et C. V. Cotton, « Explanation and Justification in Machine Learning: A Survey Or », in *IJCAI-17 workshop on explainable AI (XAI)*, 2017, vol. Vol. 8, No. 1, p. 8-13.

[4] H. Khosravi et al., « Explainable Artificial Intelligence in education », *Computers and Education: Artificial Intelligence*, vol. 3, p. 100074, 2022.

[5] M. H. Meng et al., « Adversarial Robustness of Deep Neural Networks: A Survey from a Formal Verification Perspective », *IEEE Trans. Dependable and Secure Comput.*, p. 1-1, 2022, doi: 10.1109/TDSC.2022.3179131.

[6] R. S. Baker et A. Hawn, « Algorithmic Bias in Education », *Int J Artif Intell Educ*, vol. 32, no 4, p. 1052-1092, déc. 2022, doi: 10.1007/s40593-021-00285-9.