



OFFRE D'EMPLOI



OFFRE DE THÈSE

Titre : Apport des connaissances du domaine à l'apprentissage automatisé

Institut de recherche : Laboratoire d'informatique, de modélisation et d'optimisation des systèmes (LIMOS) – École des Mines de Saint-Etienne

Candidatures à envoyer à : audrey.bruyere@vinci-energies.com



COURBON SOFTWARE
recrute :

DOCTORANT (F/H) CONTEXTE

La thèse proposée s'inscrit dans le contexte de l'Industrie du futur (ou Industrie 4.0), vouée à être plus flexible et plus autonome grâce aux technologies du numérique et de la mégadonnée (Big Data).

L'entreprise Courbon Software, qui édite des logiciels pour l'informatique industrielle, dispose aujourd'hui d'une quantité importante de données issues d'activités industrielles qu'elle souhaite valoriser en proposant de nouveaux services à ses clients. Ces nouveaux services, incluant maintenance prédictive (permettant de prévoir une casse et y remédier avant qu'elle survienne) et optimisation de l'ordonnancement de tâches (par exemple, vis-à-vis d'une offre fluctuante de la part des fournisseurs d'électricité), passent principalement par l'application de techniques connues d'apprentissage automatisé et, principalement, d'apprentissage profond (deep learning).

L'application de deep learning à des données industrielles pose cependant quelques problèmes auxquels la recherche en informatique doit répondre. En particulier, la diversité des sources de données disponibles en production (données numériques, topologiques, temporelles, structurées, ...) implique un important travail de pré-traitement qui réduit d'autant la fiabilité des modèles obtenus par rapport à des évaluations sur des données-tests. Or, la fiabilité est une nécessité pour un système industriel. Par ailleurs, les modèles obtenus avec de l'apprentissage profond sont difficilement interprétables, notamment d'un point de la responsabilité légale (pour établir qui de l'éditeur de logiciel ou de l'entreprise ayant fourni les données est responsable) lorsqu'un dysfonctionnement est observé sur une ligne de production.

Voir par exemple :

- Gusmeroli S., Dalle Carbonare D. (eds) (2020). [Big Data challenges in Smart Manufacturing Industry](#) (v. 2020). Brussels. BDVA.



COURBON SOFTWARE
recrute :

DOCTORANT (F/H) SUJET ET THÈMES ABORDÉS

L'approche proposée dans la thèse est d'intégrer des connaissances formelles à des techniques d'apprentissage profond sur des données industrielles. Les connaissances formelles peuvent être représentées sous formes de formules logiques, de données relationnelles ou, plus communément, de graphes de connaissances, à l'instar du [Google Knowledge Graph](#) ou de [Wikidata](#). Elles permettent à la fois d'intégrer dans une même base de connaissances des sources hétérogènes de données et de fournir un cadre d'"explicabilité" à des algorithmes d'apprentissage.

Les travaux effectués durant la thèse seront basés sur les progrès récents dans l'apprentissage non-supervisé appliqué aux graphes de connaissances, comme ceux d'Antoine Bordes (aujourd'hui à Facebook AI) et de Pascal Hitzler. Ils auront pour objectif la spécification d'une méthode pour intégrer des connaissances formelles génériques à des tâches d'apprentissage distinctes telles que la détection de rebuts et l'estimation d'indicateurs de performances dans l'ordonnancement de tâches.

Les thèmes abordés durant la thèse seront les suivants (liste non-exhaustive) :

- Intégration sémantique de données (RDF, SPARQL, OWL)
- Apprentissage profond (PyTorch, TensorFlow)
- Modèles de connaissances pour l'industrie (OPC-UA, AutomationML)

Voir par exemple :

- Bordes A., Usunier N., Garcia-Duran A., Weston J., Yakhnenko O. (2013). "[Translating Embeddings for Modeling Multi-relational Data](#)," Neural Information Processing Systems (NIPS). South Lake Tahoe, United States.
- Hitzler P, Bianchi F, Ebrahimia M., Sarker M.K. (2020). "[Neural-Symbolic Integration and the Semantic Web](#)," Semantic Web, vol. 11, no. 1, pp. 3-11. IOS Press.





DOCTORANT (F/H)

COMPÉTENCES REQUISES

- Master en informatique ou science des données
- Anglais courant lu
- Connaissances solides en gestion de bases de données et données massives (SQL, NoSQL, Hadoop, ...)
- Connaissances solides en programmation dans un langage haut-niveau (Python, Java, JavaScript, ...)

DESCRIPTION ENTREPRISE / LABORATOIRE

Rattachée au Groupe VINCI Energies, COURBON SOFTWARE est une entreprise dynamique et innovante, spécialiste de l'édition logiciel pour la transformation digitale.

COURBON SOFTWARE conçoit des logiciels innovants intégrables à tous les process de production. Nos suites logicielles sont élaborées en fonction des exigences de chaque secteur d'activité : industrie pharmaceutique, agroalimentaire, chimie des matériaux, manufacturier. COURBON SOFTWARE accompagne ainsi la transformation digitale des entreprises et participe à la croissance de leur performance industrielle. COURBON SOFTWARE se distingue par des logiciels innovants et totalement adaptables à tous les processus de fabrication.

Le Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes (LIMOS), qui sera le laboratoire d'accueil, est une Unité Mixte de Recherche (UMR 6158) en informatique, et plus généralement en Sciences et Technologies de l'Information et de la Communication (STIC). Le LIMOS est principalement rattaché à l'Institut des Sciences de l'Information et de leurs Interactions (INS2I) du CNRS et de façon secondaire à l'Institut des Sciences de l'Ingénierie et des Systèmes (INSIS). Il a pour tutelles académiques l'Université Clermont Auvergne et Mines Saint-Etienne (MSE). Le positionnement scientifique du LIMOS est centré autour de l'Informatique, la Modélisation et l'Optimisation des Systèmes Organisationnels et Vivants. Le poste est à pourvoir dans l'axe des systèmes d'information et de communication (SIC).