

Offre de Stage à Airbus Defence and Space (Advanced Information Processing)

Sujet de Stage : Apprentissage automatique profond pour l'extraction d'information et d'évènements à partir de texte.

Cadre du stage :

Ce stage recherche, d'une durée de 6 mois, se déroulera sur le site Airbus d'Elancourt au sein d'une équipe de Recherche & Développement spécialisée dans le traitement massif de l'information non structurée (*Big Data*). Cette équipe est impliquée dans des projets d'études amont ainsi que divers programmes de recherche partiellement financés par l'Agence Nationale de la Recherche, l'Agence de Défense Européenne ainsi que l'Union Européenne. Une poursuite du stage dans le cadre d'une convention CIFRE est envisagée sur une problématique de population de base de connaissances à partir de texte. L'encadrement sera assuré par les membres du département (Leila Khelif et Bruno Grilhers) ainsi que par des chercheurs du LIPN (Haïfa Zargayouna et Thierry Charnois)

Contexte :

Le département R&D développe actuellement, en s'appuyant sur le socle technique open source OW2 WebLab, une solution de veille nommée FORTION MediaMining. Cette solution vise à fournir une solution complète de collecte d'information multimédia –texte, image, audio, vidéo - disponible en sources ouvertes (web, réseaux sociaux), d'analyse (extraction et recherche d'information, transcription de la parole, traduction automatique, etc...) et d'exploitation (visualisation spatio-temporelle, réseau relationnel, statistiques, etc.). Celle-ci dispose notamment de fonctionnalités d'extraction d'information de relations et d'évènement à base de patrons linguistiques.

Objectifs :

Ce stage vise à étudier la possibilité de remplacer / hybrider les systèmes d'extraction à base de patrons linguistiques par des systèmes à base de réseaux de neurones profond, type réseau de neurones récurrents. Les différentes étapes du travail à réaliser sont les suivantes :

- Etat de l'art des méthodes d'extraction d'information, de relation et d'évènement à base de méthode d'apprentissage,
- Identification de solutions open source ou de laboratoire,
- Mise en œuvre des solutions techniques (apprentissage sur base annotée),
- Comparaison et évaluation des approches sur un corpus de référence

Domaines techniques/compétences informatiques :

Profil recherché :

Master 2 en Informatique(orienté recherche)

Domaines techniques :

Intelligence Artificielle, Apprentissage automatique, Traitement automatique des Langues, Fouille de données, Extraction d'information à partir de textes.

Compétences en développement logiciel :

Java, Python

Compétences considérées comme un plus :

Connaissance de Framework de Deep Learning type Keras, Tensorflow

Bon niveau en Anglais exigé.

La sélection se fera en deux temps : une pré-sélection par le LIPN suivi de la sélection finale par Airbus Group. Les locaux d'Airbus sont dans une zone à accès restrictif. Cela impose que le(a) candidat(e) soit habilité(e) (CD-Special France). La procédure d'habilitation prend au minimum deux mois, il nous est donc impossible de sélectionner des candidats étrangers.

Début souhaité : mars-avril 2018

Modalité de dépôt de candidature :

Les candidatures seront ouvertes jusqu'à sélection d'un(e) candidat(e) et au plus tard le 16 février.

Merci d'envoyer un CV détaillant la formation et l'expérience acquise, les bulletins de notes ou appréciations d'enseignants dans les compétences ciblées ainsi qu'une lettre de motivation à :
haifa.zargayouna@lipn.univ-paris13.fr, thierry.charnois@lipn.univ-paris13.fr

Lieu du stage : Airbus Defence and Space V
Advanced Information Processing, V
1, Boulevard Jean Moulin - CS 30503, 78997 Elancourt Cedex– France

Responsables (Airbus) :
Bruno GRILHERES (+33 (0)1 61 38 58 25, bruno.grilheres@airbus.com)
Leila KHELIF(+33 (0) 1 61 38 56 51, leila.khelif@airbus.com)