

Sujets de stages

Dans le cadre du projet de recherche « Check4Decision » financé par le CEA-MITIC, nous cherchons des stagiaires pour une durée de 06 mois avec :

- Niveau BAC+5 ;
- Connaissances en Python ;
- Connaissances et expériences souhaitées du Machine-Learning ;
- Connaissances et expériences des algorithmes des graphes appréciées ;
- Connaissances des technologies web sémantiques appréciées ;
- Certaines connaissances requises en fonction du sujet de stage ;
- Apte à lire et exploiter les publications scientifiques du domaine ;
- Maîtrise de l'anglais scientifique obligatoire ;
- **Disponibilité de 06 mois et résidence du stage à THIES obligatoire.**

Merci d'envoyer au plus tard le 08 Avril 2019 à l'adresse mail osall@univ-thies.sn un dossier de candidature composé de :

- Un CV ;
- Une demande de stage spécifiant le sujet du candidat ;
- Une lettre de motivation.

Seuls les dossiers retenus seront convoqués en entretien. *Rémunération à convenir lors de l'entretien.*

Sujet de stage 1 : Modélisation des données pour la visualisation en grande dimension des faits journalistiques

Il s'agit dans ce stage de mettre en place un module de visualisation basé sur le concept de « linked data » pour permettre de mettre en évidence les relations intrinsèques entre les différentes parties des articles.

Ce module a pour objectif principal de construire une « structure de données » adaptée à l'exploration des articles collectés.

Sujet de stage 2 : Conception d'un algorithme de collecte large-spectre d'articles

Dans ce stage il sera question de mettre en place un algorithme paramétrable et multi-cible pour la collecte d'articles.

L'étudiant(e) sélectionné(e) devra :

- Concevoir une interface qui intègre les différents paramètres d'une collecte d'articles,
- Mettre en place une base de données NoSQL pour le stockage des données collectées,
- Proposer un algorithme de collecte multi-événement capable de faire des collectes large-spectres sur une longue durée.

Sujet de stage 3 : Conception d'un tableau de bord pour l'exploration des faits

Le but de ce stage est de mettre en place une base de connaissances pour l'exploration des données collectées.

Le travail consistera en :

- Définir une typologie des données collectée qui se prête à un archivage « intelligent » pour en faciliter l'exploration,
- Proposer un modèle de ces données pour leur stockage à moindre coût,
- Mettre en place un moteur de recherche basé sur la suite Elasticsearch

Sujet de stage 4 : Conception d'un algorithme de classification semi-supervisée des faits journalistiques

Dans ce stage il est demandé d'explorer les possibilités qu'offre l'Intelligence Artificielle pour construire un algorithme de classification

Bien que l'approche privilégiée soit les méthodes semi-supervisées, il n'est pas exclu d'envisager d'autres approches pouvant apporter des résultats satisfaisants.

Sujet de stage 5 : Détection de mirco-expressions dans une vidéo

Ce stage sert de prétexte pour constituer une base de données d'expressions faciales sur des personnes d'origine africaine et entraîner un réseaux de neurones à convolution ou à défaut proposer selon la bibliographie une approche d'apprentissage avec une meilleure performance. Un travail a été entamé seulement le RNA construit est plus performant sur les visages blancs que noirs du fait d'un défaut d'apprentissage avec une base contenant des personnes avec des visages noirs. Ce sujet nécessite des connaissances de base en python, machine learning, une connaissance des outils d'apprentissage artificiel comme TensorFlow, Keras,...

Sujet de stage 6 : Prédiction de fact checking basé sur l'analyse bayésienne

Ce stage consiste à compléter le travail de détection de fact-checking déjà entamé dans la thèse d'Edouard et de proposer un modèle bayésien de détection de fact-checking. Il s'agira de prendre en compte les à priori et de combiner dans un likelihood la veracité des faits et la fiabilité de l'émetteur du fait pour calculer un posterior.

Sujet de stage 7 : Constitution du réseau d'un fait

Ce stage consiste à bâtir un réseau de faits et commentaires autour d'un fait. Ce réseau peut être constitué de tous les articles et de leur source qui parlent du fait,

des faits similaires à ce fait, des commentaires et appréciation sur le fait.
L'étudiant procédera à une exploration dynamique du réseau.