

Legal Document Deepfake : Etat-de-l'Art de la détection de fraude documentaire appliqué au milieu judiciaire



Description

Au sein du Pôle Recherche de l'AMIAD (Agence Ministérielle pour l'IA de Défense), vous rejoindrez notre équipe, où vous jouerez un rôle clé en explorant des approches fondées sur des données multimodales (images, texte, son).

Créée en 2024, l'Agence ministérielle pour l'intelligence artificielle de défense (AMIAD) constitue la première mesure concrète de la stratégie IA du ministère des Armées. Directement rattachée au ministre des Armées, elle a pour mission première d'équiper les forces en solutions d'intelligence artificielle, en transformant les expérimentations en capacités concrètes. Elle joue ainsi un rôle central dans le passage à l'échelle des technologies d'IA de défense. Son objectif est de renforcer la supériorité opérationnelle et l'efficacité quotidienne du ministère, en réponse aux besoins exprimés par les forces. L'AMIAD est composée de deux pôles : le pôle technique et le pôle recherche. Si le premier a vocation à concevoir et déployer des solutions d'IA destinées aux forces, qu'il met en production à grande échelle, le second vise à préparer l'avenir en anticipant les verrous scientifiques et techniques. C'est au sein du Pôle Recherche que ce stage est proposé et sera donc orienté vers des approches exploratoires et à but scientifique.

Ce stage vise à étudier, dans le cadre légal, les documents falsifiés automatiquement par des outils génératifs. Suite à l'augmentation des attaques par documents falsifiés (usurpation d'identité, prise de contrôle d'entreprises, etc.), il existe un fort besoin de détecter des fraudes documentaires (deepfake documentaire) qui peuvent revêtir plusieurs formes :

- document contrefait : fabriqué intégralement de zéro ;
- document authentique falsifié : un document réel dont on a altéré une ou plusieurs informations (date, nom, etc.) ;
- document authentique sur base frauduleuse : un document obtenu en donnant des justificatifs frauduleux.

Ce stage vise à s'attaquer aux deux premiers points et à établir un état-de-l'art des techniques et approches permettant cette détection de faux documents.

Objectifs

L'objectif est de proposer une première base de travail pour une collaboration avec l'institut Robert Badinter, groupement d'intérêt public pour les études et la recherche sur le droit et la justice. Le stage se développera selon deux axes parallèles de travail :

- Le premier axe consistera donc à établir un état de l'art des deepfakes documentaires : dresser l'inventaire des corpus et des évaluations existants, ainsi que des approches utilisées pour créer et détecter ce type de documents contrefaits. L'une des pistes pour le premier sujet sera certainement l'étude des outils permettant de générer des données contrefaites.
- Le second axe concerne la mise en place de la collaboration avec l'Institut Robert Badinter, notamment par l'organisation de réunions de travail afin de préciser les domaines, les types de documents pertinents pour les acteurs judiciaires et les professions juridiques, et de comprendre les méthodes de travail des magistrats.

Profil recherché

En formation d'ingénieur ou de master spécialité machine learning / linguistique informatique / data science et vous disposez :

- de connaissances préalables dans les domaines du traitement automatique des langues, de l'apprentissage automatique et de l'apprentissage profond ;
- d'un intérêt prononcé pour la recherche et les nouvelles technologies et êtes à l'aise en anglais ;
- de créativité et de motivation ;
- de solides compétences en programmation python et familier l'un des frameworks IA (idéalement PyTorch).

Informations générales

- Durée : 6 mois
- Niveau : Master 2 (ou équivalent en école d'ingénieur)
- Début : dès que possible
- Lieu : AMIAD - X-Novation Center, École Polytechnique, Palaiseau
- Ressources mises à disposition : l'AMIAD dispose de ressources computationnelles et dédiées (Serveurs et GPU) ainsi qu'un accès à Jean-Zay, le super-calculateur du CNRS.
- Contacts : Christophe Servan, CR AMIAD, prenom.nom@polytechnique.edu