



Rejoignez Thales, **leader mondial des technologies de sûreté et de sécurité pour les marchés de l'Aérospatial, du Transport, de la Défense et de la Sécurité**. Fort de **80 000 collaborateurs dans 68 pays**, le Groupe bénéficie d'une implantation internationale qui lui permet d'agir au plus près de ses clients, **partout dans le monde**.

Intitulé du stage :

Explication d'un algorithme d'apprentissage par renforcement

Situation organisationnelle: Thales Research & Technology

Localisation: Palaiseau (campus Ecole Polytechnique)

Description :

Les algorithmes d'apprentissage par renforcement ont un potentiel énorme dans les applications industrielles. Il s'agit typiquement d'aider un agent à prendre un ensemble de décisions de manière séquentielle pour atteindre un objectif final tout en s'adaptant à un environnement dynamique. Un exemple emblématique est le cas de deux joueurs qui jouent de manière stratégique l'un face à l'autre.

Le gros frein de l'insertion de ce type de technologie dans les systèmes réels est la difficulté pour un humain à comprendre les stratégies et choix obtenus par ces algorithmes. Il est essentiel que le système soit capable d'expliquer ces décisions. Ceci fait partie d'un sujet majeur en Intelligence Artificielle, nommé *explainable AI* (ou XAI).

Le sujet du stage porte sur l'implémentation d'une méthode d'explication pour un algorithme d'apprentissage par renforcement utilisée dans un jeu stratégique entre deux joueurs. L'explication devra faire l'interface entre un joueur humain, et un algorithme d'apprentissage par renforcement qui génère des recommandations.

Les missions :

Le stage commencera par un tour d'horizon des approches d'explication en apprentissage par renforcement. Parmi l'étendue des méthodes existantes, on s'intéressera en



particulier aux méthodes agnostiques aux modèles utilisés. Une méthode qui pourra être envisagée consiste à identifier les facteurs d'influence de la recommandation. La difficulté est de produire des explications qui éclairent le joueur soit dans la prise de décision à l'instant courant, soit dans une stratégie à plus long terme (accepter de perdre à court terme pour gagner à long terme).

Après une phase de conception, un démonstrateur devra être réalisé. Il se basera sur un application Thales dans un jeu de stratégie.

Profil souhaité :

Ingénieur ou Master Recherche M2

Compétence en Recherche Opérationnelle : Apprentissage Automatique

Outils informatiques : Python, Java sous Eclipse

Durée : 6 mois

Date de début souhaitée : Mars/Avril 2020

Entreprise :

Thales Research & Technology

Campus de Polytechnique

1, avenue Augustin Fresnel

91767 Palaiseau cedex

Tuteur :

Christophe Labreuche : christophe.labreuche@thalesgroup.com