

Stage Data Science

Entité : **AXA Assistance SA**

Direction : **AXA Assistance Digital & Services**

Département : **SolDev & Architecture**

Service : **Data Office**

Site AXA Assistance : 6 Rue André Gide, 92 320 Chatillon - France

Durée souhaitée : 6 mois

Date d'arrivée souhaitée : Mars/Avril 2019

Manager N+1 : Arnaud Py

Tuteurs : Mathieu Gilbert, Inès Bettaïeb

Contact pour candidater : raphael.denis@axa-assistance.com

Stage Data Science – Fraud Detection

Missions principales :

Au sein de la Direction Informatique Group d'AXA Assistance (Digital & Services), et plus particulièrement du service « Data Office», vous aurez pour mission d'appliquer des méthodes de data science sur différents sujets de connaissance client et d'optimisation des processus clés de l'entreprise.

A partir d'historiques de missions réalisées par AXA Assistance auprès de ses clients, d'historiques d'appels téléphoniques, vous mettrez en place différents modèles mathématiques pour répondre à la problématique d'optimisation de la détection de fraude (interne, prestataire, clients)

- Amélioration de l'outil existant par l'ajout de couches d'algorithmes de machine learning (unsupervised learning, outliers detection)
- Network-based fraud detection (graph, network modelling, page-rank algorithms...)
- Création d'un score de fraude minimisant le taux de faux positifs (supervised learning)
- Déploiement et industrialisation de la solution dans un environnement Hadoop (Oozie, Impala, Spark)

Ce stage se déroulera dans le cadre d'une approche Agile et sera l'occasion de :

- Découvrir de l'intérieur un acteur majeur d'assistance & assurance
- Développer une capacité d'analyse et de réponse technique face à des problématiques concrètes
- Mettre en pratique sur des problématiques concrètes et aux enjeux majeurs les techniques du Machine Learning
- Travailler dans un environnement Big Data complet (Hadoop, Spark, Python, R, Spotfire)

Profil recherché :

Etudiant(e) en dernière année d'école d'Ingénieurs ou en Master 2 Statistiques/Data Science.

- Modélisation statistique et modèles linéaires généralisés: excellente maîtrise.
- Machine learning (Gradient boosting/Random Forest, SVM, K-means,...) : excellente maîtrise.
- Langage ou librairie de statistiques (Python, Pandas, ScikitLearn) : Bonne maîtrise.
- Network modeling (community detection, dynamic networks...) : Bonne maîtrise
- Participation à des projets open-source, des compétitions Kaggle ou des meet-up de data science est un plus.
- La maîtrise de Spark et la connaissance du framework Hadoop sont un plus sans être un prérequis.

Compétences spécifiques attendues (informatiques, langues...) :

Anglais (écrit / lu / parlé) – niveau B2/C1

Informations complémentaires

Vous savez faire preuve des qualités suivantes :

- Dynamisme, esprit d'initiative dans un environnement peu balisé, autonomie, bonne aisance relationnelle
- Esprit analytique
- Rigueur et esprit de synthèse
- Capacité à travailler en équipe
- Sens de l'organisation