****

**Expression de besoin**

Campagne Stage 2018

**Merci de compléter l’ensemble des champs afin que la fiche de poste puisse être validée**

**Date de la demande : 13/01/2019**

**Date d’entrée souhaitée :** 04/03/2019

**Flexibilité :** [x]  **1 mois** [x]  **2 mois** [x]  **> 3 mois**

**Durée du stage : 6 mois**

**Nombre de poste à pourvoir : 1**

|  |
| --- |
| **Informations générales** |

**Entité : Crédit Agricole SA**

**Direction d’accueil** *(pas de sigle)* **: Marketing et Communication**

**Pôle** *(sigle)***: DCI**

**Maître de stage : Hugo Hamad : hugo.hamad****@credit-agricole-sa.fr**

**Type de métier :** Marketing et Communication

**Type de métier complémentaire si nécessaire** : Choisissez un élément.

**Lieu de travail :** Montrouge

**Y avait-il un autre stagiaire précédemment sur les mêmes missions ? Si oui, précisez le nom et la période de présence ?** Non

|  |
| --- |
| **Description de la mission** |

**Intitulé du poste :** Data Engineer Stagiaire – **Mise en place d’une plateforme de traitement et de calcul**

**Présentation du service :**

Au sein du Pôle Développement Clients et Innovation, le DataLab Groupe Crédit Agricole est un centre de compétences dédié aux sciences de la donnée et à leurs applications dans le domaine bancaire. Son rôle est de créer des approches innovantes pour la valorisation de la donnée interne et externe, qu’elle soit structurée ou non structurée. Dans le cadre de ses missions, des thématiques scientifiques à forte valeur ajoutée sont étudiées : Apprentissage Automatique, Auto-ML, Traitement du Langage Naturel, Process Mining, Time Series Mining, Deep Learning, Géomatique, *etc*. Ces activités sont menées conjointement avec des partenaires internes : les Caisses Régionales, les Entités du Groupe et les Producteurs Informatiques. Le DataLab développe également un réseau de partenaires externes lors de missions industrielles (Editeurs de logiciels, startup, SSII, *etc*.) ou de collaborations universitaires.

Dans le cadre de ce stage, vous rejoindrez l’équipe Data Engineering afin de contribuer à la valorisation de la Data et participer à la mise en place d’approches innovantes au service des Clients du Groupe.

**Descriptif de la mission :**

Contexte et objectifs du stage :

Les missions et projets du Datalab demandent la mise en place de moteurs de recherche, de recommandation, de reconnaissance ; la mise en place de modèles utilisant des algorithmes de machine et deep learning. Des chaînes de traitement sont mise en place pour gérer les aspects donnée et calcul. Ces chaînes doivent répondre à des problématiques de passage à l ‘échelle, de sécurité et d’exploitabilité. Elles doivent être également efficaces en termes d’utilisation de ressources (serveurs, réseau…).

Ce stage de Data Science a pour objectif de concevoir et développer et d’intégrer dans les infrastructures existantes une plate-forme de traitement facilitant le développement d’applications intensives en traitement de données et calcul (type deep learning).

A partir des projets en cours au Datalab (et de leurs infrastructures existantes) le stagiaire pourra proposer et mettre en place une plate-forme en utilisant des ressources (serveurs) internes ou cloud. Il définira et implémentera les politiques de monitoring, de partage de ressources et de capacity planning et d’exploitation.

Organisation et livrables :

Au cours de son stage, le Data Scientist aura pour mission de :

* Établir un état de l’art sur les chaînes de traitement comparables
* Mettre en place un tel système sur une chaîne deep learning
* Réaliser une étude comparative entre l’approche proposée et les approches de l’état de l’art
* Rédiger un rapport pour décrire la méthodologie proposée et les résultats obtenus

|  |
| --- |
| **Profil recherché** |

**Niveau d’études préparé :** [x] **Bac + 2** [x] **Bac + 3** [x] **Bac +4** [x] **Bac +5**

**Formation :** [x] **Université** [x] **Ecole de commerce** [x] **Ecole d’ingénieur** [x] **Autre précisez :**

**Ecoles ou formations dans lesquelles vous souhaitez éventuellement diffuser l’annonce :**

X, IMT, EPITA, ENS, Ecole des Mines, Telecom Paris, Ecole Centrale, INSA, UPMC, Université Paris Dauphine, Université de Lyon

**Spécialisation :**

Data Science, Machine Learning, Intelligence Artificielle

**Compétences techniques ou spécifiques au poste** *(5 maximum)***:**

* Architecture orientée données
* Bases de données SQL et NoSQL
* Devops
* Calcul distribué : Spark, etc.
* Développement Python, shell

**Compétences générales et transverses***(5 maximum)* **:**

* Autonomie et communication
* Rigueur
* Innovation/R&D
* Force de proposition
* Travail en équipe

**Langues 1 :** Français **Niveau requis :**Opérationnel (usage régulier)

*(Si aucune langue souhaitée indiquer N/A)*

**Langues 2 :** N/A **Niveau requis :**Choisissez un élément.

*(Si aucune langue souhaitée indiquer N/A)*

|  |
| --- |
| **ASPECTS BUDGETAIRES** |

**PARTIE A COMPLETER OBLIGATOIREMENT**

**CGB de facturation :**

 **Unité d’affectation opérationnelle (UO) :**

**Votre correspondant budgétaire**

**Avez-vous reçu l’accord de votre correspondant budgétaire ?** OUI NON

Signature et nom du Manager: Signature et nom du correspondant budgétaire :