



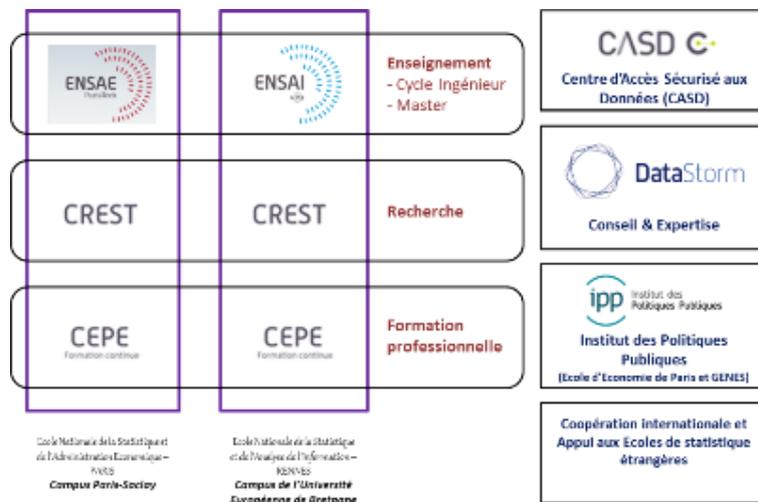
RECRUTE UN

Data Scientist Senior – développeur python

Disponibilité : dès que possible

Filiale à 100 % du Groupe ENSAE (ou GENES), entreprise de droit privé, DATASTORM a été créé en 2013 face à une demande grandissante d'expertise dans le domaine de la statistique, de l'économie, de la Data Science et de l'intelligence artificielle. Installés sur le campus de l'Ecole polytechnique du Plateau de Saclay, nous sommes au coeur d'un pôle majeur d'innovation et de recherche en intelligence artificielle.

Le Groupe ENSAE est un établissement public à caractère scientifique culturel et professionnel (EPSCP) regroupant 2 écoles d'ingénieurs (ENSAE ParisTech à Paris et ENSAI à Rennes), un centre de formation continue (CEPE, Paris), une technologie d'accès sécurisé aux données de statistique publique (CASD, Paris) et un centre de recherche en économie et statistiques (CREST, Paris et Rennes).



Avec le souhait de rester agile, DATASTORM privilégie une organisation très à plat avec un regroupement de l'ensemble des équipes opérationnelles au sein d'une direction unique pilotée par le directeur général et 2 directeurs adjoints. L'effectif, outre ces 3 directeurs, se compose à date de 5 Data Scientists Seniors et de 5 Data Scientists.

Notre culture de l'innovation se traduit directement dans notre organisation interne puisque nous avons choisi de laisser à chaque collaborateur de l'entreprise 20% de son temps pour travailler sur des sujets de R&D et de veille (développement de packages open-source / produits, participation des compétitions de data science / conférences / formation). Outre les bénéfices pour notre entreprise des produits de ces travaux de R&D, ceci offre à nos collaborateurs la possibilité de se tenir en permanence à la pointe des outils et méthodes mis en œuvre. DATASTORM dispose d'ailleurs de son propre dépôt ouvert sur GitHub (<https://github.com/datastorm-open>), avec des contributions à l'éco-système open source (packages R et python).

DATASTORM a développé un portefeuille client très diversifié au cours de ses 5 premières années d'existence, notamment dans les secteurs de l'énergie, des médias et de la banque et assurance :



Les locaux de DATASTORM se situent au sein de l'ENSAE, sur le plateau de Saclay au 5 Avenue Le Chatelier, 91120 Palaiseau. Nous disposons également de bureaux de passage à Malakoff. De plus, une partie des études se réalisent également au sein des sites de nos clients, la plupart en région parisienne.

Vos missions :

Rattaché au directeur adjoint en charge des opérations, vous interviendrez sur des projets à forte valeur ajoutée. Par exemple :

- Expérimentation R&D autour du Big Data et mise en place de modèle de Machine Learning et de Deep Learning exploitant les données détaillées des transactions bancaires
- Mise en place de modèles de prévision de production en énergie éolienne et photovoltaïque à la maille France, et intégration au SI de l'entreprise
- Réalisation d'un prototype data Analytics pour l'estimation des défauts sur les canalisations
- Développement d'un système de dédoublement des stocks par text-mining
- Mise en place d'une application pédagogique destinée à visualiser l'utilisation du réseau de transport.

En tant que Data Scientist expérimenté au sein notre équipe pluridisciplinaire vous contribuez, en collaboration avec un Data Scientist Junior et, suivant les études, avec un chercheur, aux projets de pour nos clients :

- Gestion des données (récupération, préparation, stockage des données structurées et/ou non structurées)
- Analyse du problème avec les méthodes pertinentes (classification supervisée et non-supervisée, modélisation, optimisation, système de recommandation, text-mining)
- Data Visualisation des résultats
- Développement de POC (*Proof of concept*)
- Packaging / déploiement / restitution des résultats
- Capitalisation autour des solutions mises en œuvre

De par la diversité des projets, et à travers les 20% de temps affectés à la R&D, vous monterez en compétences sur les différentes technologies autour de la Data Science.

Votre profil :

- Bac + 5 (école d'ingénieur ou université), avec une **spécialisation Statistiques/Mathématiques appliquées ou Data Science**
- 5 à 10 ans d'expérience
- Maîtrise de méthodes d'analyses de données, des modèles statistiques et de Machine Learning
- Familier du langage python en particulier des packages numpy, scikit-learn, panda, flask, dash pour ce poste très orienté Data Science. La connaissance de R est un plus
- Rigueur dans le processus de développement (GIT, tests unitaires, ...)
- Expérience de mise en production et de monitoring d'applications
- A l'écoute du monde open source académique, des nouvelles technologies et des nouveaux algorithmes
- Rigueur dans le travail
- Autonome et force de propositions

Contact

Les candidat(e)s sont invité(e)s à adresser leur candidature (CV, lettre de motivation) par mail à l'adresse suivante : contact@datastorm.fr

Des précisions sur la fiche de poste peuvent être demandées à :

- Benoît Thieurmél, Directeur des opérations
- Benoit.thieurmel@datastorm.fr
- 06 69 04 06 11