

Stagiaire Data Scientist à l'ILB Data Lab

Les ILB Labs, la cellule de R&D de l'Institut Louis Bachelier

L'ILB est un réseau de recherche partenariale en économie et en finance. Il a été créé en 2008 par le ministère des Finances pour établir un écosystème de recherche qui rassemble académiques, entreprises et pouvoirs publics. Sa mission est d'amplifier l'impact de la recherche et de la modélisation dans ces domaines.

Les ILB Labs sont une **équipe mixte** d'ingénieur.e.s de R&D spécialisé.e.s en data science (DATA LAB) et d'analystes en finance durable (ESG LAB). Son objectif est d'accompagner les entreprises dans leurs transformations **numérique** (analyses statistiques et développement de modèles machine / deep learning pour répondre à une problématique métier) et **environnementale** (alignement climatique d'activité, notation verte, risque climatique, ingénierie financière verte, estimation d'impact de projets...). L'équipe mène **principalement des projets dans le secteur financier** (avec des acteurs de type banque, assurance, gestionnaire d'actifs...).

L'équipe bénéficie d'une expertise unique grâce au réseau de plus de 1000 chercheurs de l'ILB, parmi lesquels elle peut solliciter des experts à la pointe de la recherche sur des thématiques précises afin de l'accompagner dans ses projets, que ce soit en data science, finance durable ou de nombreux autres domaines (immobilier, sécurité...).

Le positionnement des ILB Labs est ainsi à **mi-chemin entre recherche et conseil** : l'objectif est de fournir des livrables opérationnels de pointe dans le cadre de projets encadrés par des experts académiques de haut niveau, dans des formats de gestion de projets fortement inspirés du conseil.

Précisons maintenant le champ d'action du Lab pour lequel tu postules : [le DATA LAB](#) !

Les projets du Data Lab

L'ensemble de l'équipe travaille sur des projets de R&D en data science en partenariat avec des acteurs de l'écosystème de la finance et de l'assurance. Nous travaillons par exemple sur :

- L'utilisation de données externes (open banking, crédit bureau...) pour améliorer les modèles de risque de crédit via ML ;
- L'extraction de données financières et extra financières des rapports d'entreprises via NLP afin de créer de nouveaux indicateurs ESG ;
- L'application de méthodes de ML pour prédire la sinistralité d'un portefeuille d'assurances ;
- La revue de l'état de l'art des méthodes existantes en Graph Neural Networks pour les appliquer au cas d'usage de détection de fraudes, lutte anti-blanchiment et financement du terrorisme ;
- Le développement d'un indice statistique pour classer les places financières mondiales, dont la méthodologie a pour vocation d'être explicite et publique...

Selon les projets, l'équipe est en interaction avec des académiques et expert.e.s renommé.e.s sur les thématiques abordées.

Les objectifs du stage

Ce stage pourra s'articuler autour de diverses thématiques, selon les projets qui commenceront ou seront au cours au moment de ton arrivée. Des exemples sont donnés à titre illustratif dans le paragraphe précédent.

Tu auras un projet principal et pourras être affecté.e sur des projets plus courts. Comme en conseil, il nous est difficile de savoir en avance les types de projets qui seront lancés et de connaître en avance la technicité du sujet.

Dans ce cadre et en lien direct avec les responsables des ILB Labs, tu auras pour mission de :

- Comprendre, analyser et préparer les données de façon justifiée et adaptée au contexte du projet (données, contraintes, besoins...) ;
- Identifier les algorithmes adéquats, potentiellement à la suite d'une revue de littérature ;
- Mettre en place l'approche de modélisation et analyser la robustesse / les possibles biais des algorithmes, en appliquant les bonnes pratiques de code suivies dans l'équipe ;
- Être capable d'interpréter les résultats obtenus en prenant du recul ;

- Choisir et développer des éléments de visualisation de qualité permettant de bien illustrer les messages que l'on souhaite faire passer ;
- Savoir justifier les choix / hypothèses pris dans le cadre du projet.

L'équipe

Tu rejoindras une équipe dynamique d'une vingtaine de personnes issues de formations d'excellence (grandes écoles, université...), avec des profils internationaux. Une forte cohésion existe au sein des Labs et c'est une qualité que nous souhaitons maintenir tout au long de notre croissance.

L'ILB est profondément engagé envers le bien-être de ses employé.e.s. En encourageant l'équilibre entre la vie professionnelle et personnelle et en offrant des opportunités de développement, nous proposons un environnement où chaque personne peut pleinement s'épanouir.

Nous encourageons les candidatures de personnes de divers horizons et expériences, même si elles ne possèdent pas toutes les compétences mentionnées. Nous croyons que la diversité des perspectives et des expériences est essentielle à notre succès en tant qu'entreprise.

Le profil recherché

Nous recherchons une personne ayant les compétences suivantes :

Compétences techniques

- Une formation de haut niveau en statistiques ou data science (école d'ingénieur, université), formation initiale ou continue, ayant permis d'acquérir des bases solides en mathématiques, sur les algorithmes de Machine Learning et les concepts statistiques fondamentaux ;
- Maîtrise de Python et des packages de data science associés (Pandas, Scikit-Learn, Matplotlib, Seaborn...) ;
- Capacité à comprendre une problématique métier et à la traduire en termes de données ;
- A l'aise en français et en anglais (écrit et oral) ;
- Maîtrise du pack Office (Excel, Powerpoint, Word), avec capacité de présenter des travaux techniques de façon précise ou synthétisée et vulgarisée ;
- Maîtrise appréciée de logiciels ou langages de gestion de bases de données (SQL...), d'outils de BI (Power BI, Tableau Desktop...) ou compréhension fine des méthodes de Deep Learning (théorique et mise en pratique).

Compétences humaines

- **Curieux.se, autonome et force de proposition** sur tous les aspects du projet, que ce soit d'un point de vue technique ou organisationnel ;
- Organisé.e, créatif.ve, à l'esprit entrepreneurial, avec de bonnes qualités relationnelles et appréciant le travail en équipe.

N'hésite pas à candidater même si tu ne coches pas toutes les cases :-) !

Le poste est localisé dans le Palais Brongniart à Paris (75002) avec télétravail à temps partiel possible (tout en respectant les conditions des écoles). La rémunération se fait selon une grille (césure 1400€ brut mensuel / fin d'études 1600€ brut mensuel) avec des tickets restaurant de 11€ par jour (4.5€ à charge de l'employé.e) et au remboursement de 50% du titre de transport sur présentation d'un justificatif. Un jour d'absence autorisée est obtenu par mois de stage.

Le processus de recrutement se déroule en 4 étapes : (Pouvant durer jusqu'à 3 semaines au total)

- Un premier entretien avec un.e Data Scientist du Data Lab pour comprendre tes motivations et apprendre à te connaître. C'est aussi l'occasion d'évaluer tes connaissances en Python et en data, avec un rapide test en live.
- Un cas pratique à réaliser à la maison en un temps limité, qui concernera une problématique classique de modélisation, afin de te mettre en situation sur un projet concret.
- Un deuxième entretien avec Marine Neyret, la responsable du Data Lab, pour discuter de ton parcours, tes motivations et pour t'expliquer plus précisément ce que l'on attend d'un.e stagiaire au Data Lab.
- Un dernier entretien avec Louis Boulanger, le directeur des Labs.

Nous restons à ta disposition pour en discuter plus en détails.

Merci d'envoyer ton **CV avec un paragraphe expliquant tes motivations, le type de stage souhaité (césure / fin d'études) et en indiquant la date à partir de laquelle tu es disponible** à : recrutement.data@institutlouisbachelier.org