



SOLUTION D'INTELLIGENCE TECHNOLOGIQUE

Fiche de poste

“DATA SCIENTIST, Machine Learning/NLP/Data Analytics”

Présentation de la société :

GraphMyTech (GMT) est une startup récemment créée qui développe et propose des solutions d'aide à la décision liée aux problématiques d'innovation et de R&D. L'analyse du portefeuille de brevets d'une entreprise et plus largement de sa base de connaissances nous permet entre autres : d'évaluer son positionnement technologique, de scénariser son développement technique, de mesurer la cohérence d'acquisition d'entreprises ou de portefeuilles de brevets, d'identifier des partenaires ayant des compétences scientifiques et techniques complémentaires, etc.

Mission :

Vous travaillerez en **étroite collaboration avec le CTO et le CEO** au développement d'outils et d'algorithmes sur la base de prototypes existants ou que vous serez amené à imaginer. Il s'agit principalement d'algorithmes de classification de documents (**NLP**) et d'analyse Big Data utilisant la **théorie des graphes**.

Vous contribuerez notamment aux trois projets en cours :

- GeT-IT : financé par la BPI
- MémoGraph : financé par la SATT Lutech.
- TIPS : financé par la Direction Général de l'Armement.

La structure est amenée à s'agrandir rapidement et il s'agit d'une bonne opportunité pour un candidat sérieux et ambitieux souhaitant se voir confier **des responsabilités croissantes** à court et moyen terme. **Possibilité de télétravail étendu.**

Profil recherché :

Indispensable :

- De formation supérieure **BAC+5 ou doctorat** avec un profil scientifique.
- Au moins **une expérience significative** comme data scientist (salarié, stagiaire ou alternant), avec utilisation de technologies de Machine Learning/Deep Learning.
- Capacité et appétence à **travailler en autonomie** dans une petite équipe.
- Une expérience en **NLP** (Natural Language processing), particulièrement en classification et clustering de documents.
- Très bonne maîtrise de **Python et Sklearn**
- Bonne connaissance de l'écosystème **Keras/Tensorflow** ou de **Pytorch**
- Bon niveau d'anglais

Souhaitable :

- Compétences DevOps/SysOps, bonne maîtrise de **Linux** et de l'administration serveur
- Connaissances en théorie des graphes
- Une expérience Google Cloud Engine (**GCE**) ou Amazon Web Services (**AWS**)
- Une expérience Hadoop/**Spark**
- Anglais courant