

Compétences pour être un Data Scientiste

Un data scientist est un professionnel qui est capable de traiter les données et aidés le business dans les prises de décision au quotient pour son bon fonctionnement. A travers les modèles de Machine Learning, Deep Learning.

Compétences techniques (soft skills)

1. **Curiosité** : la data science est un domaine qui change et qui évolue très vite donc être curieux nous permet d'apprendre et d'être informé sur les nouvelles technologies pour être toujours au top.
2. **Travail en équipe** : comme nous travaillons avec plusieurs corps de métiers donc avoir des compétences relationnelles est très importants.
3. **Communication** : Savoir communiquer de façon simple est primordial car nous travaillons avec des gens parfois qui n'ont aucune en data science.
4. **Organisation** : la capacité d'organisation est essentielle car nous travaillons dans un domaine très vaste donc savoir s'organiser d'être productif.

Compétences comportementales (hard skills)

1. Mathématiques :

Probabilité : comprendre et savoir interpréter les probabilités pour pouvoir bien manipuler les données.

Algèbre linéaire : comprendre et savoir manipuler les espaces vectorielles c'est-à-dire les vecteurs et les matrices car les données sont représentées sous formes vectorielles.

Optimisation : apprendre et approfondir tes connaissances en optimisation mathématiques.

Analyse : comprendre et savoir interpréter les fonctions dérivées et intégrales en mathématiques.

2. Statistique :

Statistique descriptive : analyse uni et multivariée, comprendre et savoir interpréter les variables quantitatives et qualitatives (outil statistique en fonction du type de variable).

Statistique inférentielle : comprendre la corrélation entre les variables, etc.

3. Types d'apprentissage :

Connaitre les différents types d'apprentissages et quelques algorithmes et savoir utiliser un algorithme approprié à chaque de type de données et comment utiliser les hyper paramètres.

4. Deep Learning :

Comprendre le Deep Learning pour résoudre les problèmes plus complexes.

5. Algorithmique

Connaitre l'algorithmique pour pouvoir résoudre des problèmes de vie réelle en langage machine car ce sont des machines qu'on utilise pour la data science.

6. Connaitre le SQL et NoSQL

Connaitre SQL et NoSQL pour le stockage et manipulation des données est très important.

7. Business

En général nous travaillons avec les gens du métier donc nous devons connaître les fondements du métier pour lequel travaillons.