

DATA SCIENTIST / INGÉNIEUR.E BASE DE DONNÉES

Mission générale

Valorise l'ensemble des données de l'entreprise pour l'aider dans sa prise de décisions et en faire un levier de création de valeur.

Analyse des masses de données hétérogènes, éventuellement non structurées, pour en extraire de la connaissance utile à l'optimisation des offres et services de l'entreprise.

Possède une vision transverse et croise les données de différentes sources dispersées.



ACTIVITÉS PRINCIPALES

> Extraction, uniformisation et structuration des données

- Collecte, sélection et validation des données pertinentes
- Définition des solutions de stockage et structuration des données
- Conversion, codage et cartographie des données
- Détermination des outils, méthodes d'acquisition de données
- Conception de l'architecture d'un entrepôt de données décisionnelles (Data Warehouse)

> Analyses prédictives et développement de la connaissance client

- Modélisation statistique des données
- Algorithmes d'apprentissage et scénarios prédictifs
- Optimiser la segmentation client à l'aide des statistiques
- Modèles détection insights1 et opportunités de marché

> Développement d'outils de support aux clients internes

- Participation à la stratégie marketing de l'entreprise
- Analyse de données pour systèmes d'aide à la décision
- Participation au développement des indicateurs de performance de l'entreprise
- Réalisation et présentation d'études statistiques pour les clients internes
- Animation des ateliers d'expression des besoins internes
- Outils de reporting dynamique et multidimensionnel (OLAP4)
- Formation des utilisateurs aux outils informatiques et décisionnels

> Veille technologique sur les outils de datamining – Archivage

> Management d'équipe interne / externe – Gestion de projets



COMPÉTENCES REQUISES

> Maîtrise de l'anglais technique et scientifique oral et écrit

> Transverses

- Rigueur, concentration, méthode
- Sens de la communication, de l'écoute, pédagogie, force de conviction et de proposition
- Curiosité, appétence pour apprendre

> Métier

- Maîtriser les algorithmes d'apprentissage automatique (Machine Learning)
- Maîtriser les outils de data management (SAS, SPSS, SAP Infinite Insight, Python, R, Excel, Access...)
- Maîtriser les technologies HADOOP
- Maîtriser des bases de données SQL et no-SQL.
- Avoir une bonne connaissance des réseaux de neurones et d'intelligence artificielle
- Avoir une bonne connaissance des outils de Web analyse (Omniture, Google analytics etc.)
- Avoir de solides connaissances en marketing



DIPLÔMES REQUIS

- Formation de niveau Bac +5 : Master en économétrie, informatique ou statistiques
- Doctorat en informatique, statistiques, mathématiques ou modélisation de données



EXPÉRIENCE

Profils confirmés, première expérience de 3 ans à minima.



MOBILITÉ PROFESSIONNELLE

> Au sein de la filière métier :

- Data manager

> Hors filière métier :

- Chargé.e / responsable marketing
- Chargé.e pharmaco-économie



TENDANCES D'ÉVOLUTION DU MÉTIER

L'augmentation de la quantité et de la complexité des données entraîne un besoin en profils polyvalents possédant plusieurs compétences. Cette croissance de données représente un potentiel commercial à exploiter.

Ce métier tend à collaborer de plus en plus étroitement avec le marketing. Les profils Data Scientist économétrie peuvent évoluer vers cette famille métier. Compte tenu de la forte spécialisation nécessaire pour exercer ce métier, les opportunités d'embauche sont nombreuses et supérieures à la quantité de profils qualifiés.

L'automatisation progressive des tâches de l'ingénieur base de données aura pour conséquence d'augmenter sa productivité.



TÉMOIGNAGE

Erwan

DATA SCIENTIST

PARCOURS :

après un cursus à l'ESME Sudria, marqué par un prix lors du Symposium pour un projet associant les données issues de l'assurance maladie avec la météo, Erwan a choisi de débiter sa carrière de Data Scientist dans l'industrie pharmaceutique.

Il a développé un outil capable d'observer sur les réseaux sociaux les tendances autour des effets indésirables des médicaments.

« Il faut savoir que les entreprises pharmaceutiques délivrant des médicaments doivent chercher par elles-mêmes et par tous les moyens si les patients subissent des effets secondaires. Cela fait partie de leurs obligations. Si, dans le futur, la réglementation concernant la prise en compte des réseaux sociaux se fait plus stricte, l'outil développé leur permettra d'y faire face.

L'outil que j'ai mis au point permet d'extraire les informations d'une source prédéfinie en utilisant des noms de médicaments préalablement rentrés. Par exemple, si on rentre le mot clé « aspirine », l'outil sélectionnera automatiquement tous les textes le contenant et cherchera ensuite dans ces textes les mentions faites sur les effets indésirables en se basant sur les dictionnaires médicaux – bien souvent disponibles en open source – et les notices des médicaments concernés.

Nous avons effectivement pu aboutir à des conclusions très intéressantes. L'une d'elles porte ainsi sur la fréquence des effets indésirables, qui n'est pas du tout la même sur les réseaux sociaux et dans la littérature. Par exemple, pour l'ibuprofène, si on regarde les notices d'utilisation, on apprend que l'on a plus de chance de connaître tel effet secondaire plutôt qu'un autre. En l'occurrence, il est détaillé que la somnolence est un effet secondaire assez rare. Pour autant, avec notre outil, il est ressorti que la somnolence ressortait énormément sur Twitter, faisant de cet effet celui le plus retrouvé.

L'outil nous a également permis de déceler des cas d'effets secondaires pourtant jugés comme vraiment très rares sur les réseaux sociaux. De facto, on peut également imaginer trouver de nouveaux effets secondaires grâce à lui. »