

Programme mastère Data scientist

Titre de la formation :

Expert en modélisation et apprentissage

Niveau académique :

(BAC+5)

Durée de la formation :

Généralement d'une durée d'un à deux ans

Objectif :

Former des Data Scientists hautement qualifiés capables de maîtriser toutes les étapes du cycle de vie des modèles de Machine Learning, de la préparation des données au déploiement en production, en utilisant les dernières avancées en IA, NLP, Computer Vision, et Big Data.

Description générale de la formation :

Cette formation complète couvre l'ensemble du cycle de vie des modèles de Machine Learning, du prétraitement des données au déploiement en production, en passant par l'apprentissage supervisé et non supervisé, le NLP, la Computer Vision, le déploiement en Big Data et sur le cloud.

Prérequis :

- Être titulaire d'une certification professionnelle de niveau 6 en mathématiques, informatique, technologie, économie, finance ou analyse de données ;
- Ou, être titulaire d'une certification professionnelle de niveau 5 avec une expérience professionnelle d'au moins 2 ans en mathématiques, informatique, technologie, économie, finance ou analyse de données.

Tout autre profil relève d'une procédure dérogatoire (analyse de dossier et test de positionnement).

Débouchés :

Les perspectives professionnelles des data Scientist sont diverses et variées, qu'ils choisissent de se lancer en tant qu'indépendants ou au sein des cabinets de conseil. Ces experts sont particulièrement convoités dans des secteurs tels que la finance, l'assurance, les entreprises

numériques, les télécommunications, la gestion logistique, et les grandes entreprises désireuses d'améliorer divers aspects de leurs opérations.

Options de carrière en Data Science :

L'expert en Data Science peut exercer plusieurs fonctions. Ci-après une liste non exhaustive : Data Scientist, Responsable Data, Chef de Projet Data, Business Analyst Senior, Ingénieur Machine Learning, Ingénieur Deep Learning, Ingénieur Vision Assistée par Ordinateur, Ingénieur NLP, Big Data Engineer, Chief Data Officer.

Objectifs

- Acquérir des compétences avancées dans la collecte, la préparation, la visualisation et l'analyse de données, tant structurées que non structurées, pour résoudre des problèmes métier complexes.
- Devenir un expert dans la création et l'entraînement de modèles d'apprentissage automatique supervisés pour des analyses prédictives précises, ainsi que de modèles non supervisés pour la segmentation et la réduction de données.
- Manipulation de Données Non Structurées : Développer des compétences avancées en prétraitement et analyse de données non structurées, y compris le texte et les images, afin de créer des jeux de données exploitables pour des applications spécifiques.
- Communication et Déploiement : Apprendre à présenter efficacement les résultats des analyses et des modèles d'apprentissage automatique aux utilisateurs finaux, ainsi qu'à déployer ces modèles dans des environnements opérationnels.
- Comprendre comment déployer des modèles d'apprentissage automatique à grande échelle en utilisant des technologies du Big Data pour traiter des volumes massifs de données.
- Sensibilisation aux implications éthiques de l'analyse de données et de l'IA, ainsi qu'à la prise de décision automatisée, en veillant à l'intégrité et à la transparence des processus.
- Appliquer ces compétences dans le cadre de projets concrets et d'études de cas réels, afin de développer une expérience pratique et de résoudre des défis réels liés à l'analyse de données.

Étapes d'admission

1. Inscription via un formulaire de candidature / candidature en ligne
2. Tests de positionnement en ligne : tests techniques

3. Entretien individuel de motivation
4. Pour les alternants : Montage du dossier de financement - Services d'accompagnement pour la recherche d'entreprises

Formation data scientist en alternance

- Signature du contrat d'alternance
- 18 à 24 mois en alternance : Deux modalités peuvent être proposées : 3 à 4 jours en entreprise/un jour à l'école en télé-présentiel ou un rythme de 3 semaines en entreprise et 1 semaine à l'école

Programme

Blocs de compétences	Modules	ECTS*	Volume horaire (Heure)	Semestre
Collecter, pré-traiter, analyser des données et créer des variables pertinentes (features engineering) pour l'entraînement d'un modèle	Introduction au Prétraitement des Données (Python et Pandas)	2	20	1
	Nettoyage de données	2	20	1
	Transformation et agrégation des données (Features engineering)	2	30	
	Analyse statistique et visualisation	2	20	1
Concevoir des modèles d'apprentissage à partir de données structurées et de données non structurées (texte et images)	Introduction aux Modèles d'Apprentissage Automatique Supervisé	3	30	1
	Choix des modèles et entraînement	3	30	1
	Evaluation des modèles	3	30	1
	Optimisation des hyperparamètres et validation	3	30	1
	Modèles de clustering	2	30	1
	Réduction de dimensionnalité / Features selection	2	30	1
	Fondamentaux du NLP	2	30	2
	Prétraitement du texte et analyses avancées du NLP	3	30	2
	Traitement de Langage Naturel Profond	3	30	2
	Fondamentaux du computer Vision	2	30	2
	Prétraitement d'Images et techniques avancées	3	30	2
	Deep learning en computer Vision	3	30	2

Gérer le cycle de vie et le déploiement en production d'un modèle (MLOps) ou d'une application dans des environnements classiques et Big Data	Introduction au Déploiement de Modèles ML et à l'Environnement de Développement	3	30	2
	Gestion des Modèles avec MLflow	3	30	2
	Création d'une API pour le Déploiement de Modèles	3	30	2
	Déploiement sur le Cloud et Mise à l'Échelle	3	30	2
Manager des projets IA et de data science complexes en intégrant les contraintes légales et les valeurs d'éthique	Gestion de Projets IA et Data Science (SCRUM)	3	30	1
Total		60 ETCS	600	2

Modalités d'évaluation

Bloc de compétences	Modalité d'évaluation
Collecter, pré-traiter, analyser des données et créer des variables pertinentes (feature engineering) pour l'entraînement d'un modèle	Mise en situation professionnelle (réelle ou reconstituée)
Concevoir des modèles d'apprentissage à partir de données structurées et de données non structurées (texte et images)	Mise en situation professionnelle (réelle ou reconstituée)
Gérer le cycle de vie et le déploiement en production d'un modèle (MLOps) ou d'une application dans des environnements classiques et Big Data	Mise en situation professionnelle (réelle ou reconstituée)
Manager des projets IA et de data science complexes en intégrant les contraintes légales et les valeurs d'éthique	Cas pratiques ou Evaluation écrite

Indicateurs clés de la formation Data science

- ⇒ **Taux de réussite des diplômes : 100 % diplômés** – Tous les apprenants ont obtenu leur diplôme.
- ⇒ **Taux de poursuite d'études : 0 %** – Aucun diplômé ne poursuit d'études supplémentaires.
- ⇒ **Taux d'interruption en cours de formation : 0 %** – Aucun apprenant n'a interrompu sa formation.
- ⇒ **Taux d'insertion professionnelle : 100 % en emploi** – Tous les diplômés ont trouvé un emploi.
- ⇒ **Taux de satisfaction des apprenants : 92 % satisfaits** – La grande majorité des apprenants sont satisfaits de leur expérience.
- ⇒ **Taux de recommandation (Net Promoter Score ou % de recommandation) : 90 %** – 90 % des apprenants recommandent la formation à un proche ou à un collègue.
- ⇒ **Taux d'assiduité : 98 %** – Les apprenants ont été très réguliers, avec 98 % de présence aux cours et projets de la formation Data Science.