

Préparation et intégration des données pour l'IA

RS5340 · RS · Publiée

L'essentiel

Code(s) NSF	326t, 326
Formacode(s)	31025, 31052, 31009, 30812, 30887
Type d'enregistrement	Enregistrement sur demande
Date d'enregistrement	20/01/2021
Date d'échéance	19/08/2025

Certificateur(s)

DIGITAL SERVICES SCHOOL · SIRET : 38183776400095<https://ecoledeladataetia.com>**ORANGE · SIRET : 38012986648625****BUSINESS ET DECISION UNIVERSITY · SIRET : 38183776400087****ORANGE · SIRET : 38012986646850**

Résumé de la certification

OBJECTIFS ET CONTEXTE

Depuis les cinq dernières années, les projets Data Science et d'intelligence artificielle connaissent une forte croissance dans des domaines très variés, tels que l'industrie pétrolière, la téléphonie, la distribution, la banque, la finance et les services. Les récentes évolutions technologiques permettent de faire évoluer ces projets vers la mise en place de véritables solutions d'intelligence artificielle. La réalisation de ces projets nécessite de former des équipes de « Data-Engineers » pour lesquelles il est nécessaire de conjuguer l'acquisition des compétences techniques et méthodologiques, l'expérience des cas concrets et surtout, un parcours d'apprentissage jalonné.

Les principales compétences visées par cette certification sont celles nécessaires à la mise en œuvre des activités suivantes :

- Auditer la qualité des données
- Préparer, nettoyer et programmer le recodage des données
- Programmer les indicateurs nécessaires au fonctionnement des algorithmes

COMPÉTENCES ATTESTÉES

1. Maîtriser la statistique descriptive monovariée
2. Savoir utiliser la librairie Numpy en langage Python
3. Savoir utiliser la librairie Pandas en langage Python
4. Savoir réaliser un audit de qualité de données
5. Connaître les principales règles de RGPD
6. Maîtriser la statistique bivariée
7. Savoir tester et mesurer la paramétrie des données
8. Savoir programmer en Python
9. Savoir programmer le recodage des données pour les rendre paramétriques
10. Savoir caractériser et traiter les valeurs manquantes
11. Savoir identifier et traiter les valeurs aberrantes
12. Maîtriser le fonctionnement des bases de données SQL
13. Pouvoir créer des jeux de données de test
14. Savoir mettre en œuvre une stratégie inférentielle
15. Maîtriser les bases de la statistique multivariée

16. Être capable de mettre en œuvre du machine-learning non supervisé
17. Être capable de mettre en œuvre du machine-learning supervisé
18. Savoir utiliser les algorithmes du domaine public pour créer un modèle

Voies d'accès

MODALITÉS D'ACCÈS

Formation continue

COMPOSITION DU JURY — FORMATION CONTINUE

Le jury est présidé par un data-scientist confirmé. Il établit la composition du jury qui sera composée d'au minimum 4 consultants-experts des domaines couverts par le certificat et du directeur de Business & Decision University. Les membres du jury peuvent être des formateurs du certificat, mais dans tous les cas plus de 50% des membres du jury seront des personnes extérieures à l'organisation et si des formateurs participent au jury, leur voix sera strictement consultative

Base légale

Date de décision	20/01/2021
Durée de l'enregistrement	5 ans
Date d'échéance de l'enregistrement	19/08/2025

Pour plus d'informations

Fiche officielle (France Compétences)

<https://www.francecompetences.fr/recherche/rs/5340/>