PROGRAMME DE FORMATION

Machine learning avec Python : les fondamentaux

Développement & Programmation

RÉFÉRENCE	VERSION	ÉDITION	DOMAINE
WF-0439	V2026-005	2 novembre 2026	Développement & Programmation

DURÉE	MODALITÉ	PROGRAMME	SANCTION
35 h	Distanciel asynchrone (e-learning)	10 modules 41 leçons	Attestation et certificat de réalisation

Objectifs pédagogiques

- Charger, nettoyer, explorer et visualiser un jeu de données tabulaires avec pandas, NumPy et Matplotlib.
- Formuler un problème d'apprentissage, découper les données et construire un pipeline de préparation sans fuite d'information.
- Entraîner et comparer des modèles de régression, de classification et d'ensemble avec scikit-learn.
- Évaluer un modèle avec la métrique adaptée, la validation croisée et le réglage des hyperparamètres.
- Explorer des données sans étiquette, puis sauvegarder, documenter et présenter un modèle avec ses biais et ses limites.

L'essentiel de la formation

Public visé

Développeurs Python et analystes de données qui veulent construire, évaluer et expliquer leurs premiers modèles d'apprentissage automatique avec scikit-learn, sur des données tabulaires.

Prérequis

Programmer en Python (listes, dictionnaires, fonctions, lecture d'un fichier CSV, environnement virtuel) ; à défaut, les leçons de Python (WF-0110) sont réemployées en rappel.

Connaître les notions de base de statistique descriptive (moyenne, médiane, écart type, pourcentage).

Disposer d'un ordinateur avec une connexion internet et un navigateur récent.

Modalités et déroulement

Formation à distance, en ligne, à rythme libre, organisée autour d'organisations fictives (Kestrel Mobilité, vélos électriques en libre-service avec application et API ; Imprimerie Brasselin, PME de 45 salariés sur deux sites ; Collectif Ombelle, studio de design et de vidéo ; Médiathèques de Valtorne, réseau intercommunal avec portail en ligne ; Les Vergers de Lirac, coopérative fruitière — organisations fictives, données fictives). Un test de positionnement ouvre le parcours. Chaque leçon associe une vidéo de démonstration, une fiche pratique à conserver et deux mises en pratique corrigées (code écrit dans des carnets de l'environnement d'exercice et vérifié par des tests automatiques, analyses d'erreurs, décisions de modélisation, écrits de restitution) ; chaque module se clôt par un atelier de synthèse corrigé, une vidéo récapitulative et un quiz. Les manipulations se font dans un environnement d'exercice en ligne fourni par l'organisme (terminal, éditeur de code et services simulés, signalés comme tels) pour Python, pandas, scikit-learn et Matplotlib dans des carnets Jupyter : aucune installation, aucune licence ni aucun compte payant n'est nécessaire ; l'apprenant peut ensuite refaire les exercices sur son propre poste. Temps de formation réparti à environ 60 % de pratique, 30 % de démonstrations et d'apports, 10 % d'évaluation. Accompagnement par un formateur par messagerie (réponse sous 48 h ouvrées), correction individuelle des livrables sur demande.

Méthodes pédagogiques

- Positionnement en début de parcours
- Apports en vidéo : explications et démonstrations
- Exercices d'application et mises en pratique corrigées
- Production de livrables professionnels
- Atelier de synthèse en fin de module
- Quiz et QCM d'auto-évaluation
- Accompagnement et retours du formateur

Moyens pédagogiques et techniques

- Plateforme de formation en ligne, accès individuel à distance
- Vidéos pédagogiques
- Fiches pratiques à conserver
- Modèles et supports téléchargeables
- Quiz et exercices interactifs en ligne
- Correction des productions sur grille de critères
- Messagerie avec le formateur (réponse sous 48 h ouvrées)

Modalités d'évaluation

Positionnement initial (non éliminatoire) ; correction automatique ou sur grille critériée de chaque exercice ; atelier de synthèse corrigé et quiz à la fin de chaque module ; évaluation finale interne : projet d'apprentissage supervisé sur un jeu fictif des Vergers de Lirac non vu en cours (exploration, pipeline de préparation, comparaison de modèles par validation croisée, choix de la métrique et du seuil, analyse par segment, fiche du modèle), corrigé sur grille, et questionnaire de 20 questions. Seuil de réussite interne : 70 % des critères. Une attestation de fin de formation précise les compétences évaluées ; aucune certification, ni d'éditeur ni d'État, n'est délivrée ni préparée.

Sanction

Une attestation de fin de formation et un certificat de réalisation sont remis au participant sous réserve de la réalisation du parcours.

Cette formation n'est pas présentée comme préparant à une certification enregistrée : elle ne comprend ni inscription ni présentation à des épreuves de certification, et n'ouvre droit à aucune délivrance de certification ni au financement par le CPF.

Modalités et délais d'accès

La demande d'inscription peut être effectuée par téléphone, par courrier électronique ou au moyen du formulaire en ligne.

Un entretien ou un questionnaire de positionnement permet de vérifier l'adéquation de la formation avec les besoins, le niveau initial et les prérequis du candidat.

L'inscription devient définitive après validation du dossier administratif, acceptation des documents contractuels et, le cas échéant, confirmation du financement.

Les accès à la plateforme sont généralement activés sous 48 à 72 heures ouvrées après validation complète du dossier.

Les formations à distance sont proposées en entrées et sorties permanentes, sous réserve des délais ou dates imposés par le financeur ou le certificateur.

Accessibilité aux personnes en situation de handicap

Vidéos sous-titrées et transcrites, fiches pratiques, énoncés et corrigés en version texte lisible par les lecteurs d'écran, code et commandes fournis en texte copiable (jamais en image), contrastes et tailles de police réglables dans l'environnement d'exercice, navigation au clavier, rythme libre et durée allongeable pour les exercices chronométrés. Référent handicap joignable à handicap@will-formation.com pour étudier toute adaptation (durée, supports, outils, modalités d'évaluation).

Tarif de la formation

Prix : **2 770 €**

Financement : prise en charge possible par les organismes financeurs, selon les critères d'éligibilité applicables.

Qualiopi processus certifié

La certification qualité a été délivrée au titre de la catégorie d'action suivante : ACTIONS DE FORMATION. Certificat QUA25070009. WILL FORMATION — Société par actions simplifiée à associé unique au capital de 30 000 € — NDA 11757059675 (cet enregistrement ne vaut pas agrément de l'État) — SIRET 893 713 180 00014 — 01 85 09 25 99 — info@will-formation.com — www.will-formation.com

Programme détaillé

10 modules, 41 leçons — durée totale 35 h.

01 Du Python aux données

1. Positionnement — Python et statistique descriptive
2. Rappels Python pour le machine learning
3. Calculer avec NumPy
4. Charger et manipuler des tables avec pandas
5. Atelier de synthèse — Du Python aux données

02 Explorer et visualiser les données

1. Décrire, agréger et fusionner avec pandas
2. Repérer valeurs manquantes, aberrantes et doublons
3. Visualiser pour comprendre
4. Atelier de synthèse — Explorer et visualiser les données

03 Comprendre le machine learning

1. Qu'est-ce que le machine learning ?
2. Données d'apprentissage et de test
3. Un premier modèle et sa référence
4. Atelier de synthèse — Comprendre le machine learning

04 Préparer les données pour l'apprentissage

1. Valeurs manquantes et variables catégorielles
2. Mettre à l'échelle et créer des variables
3. Assembler la préparation dans un pipeline
4. Atelier de synthèse — Préparer les données pour l'apprentissage

05 Prédire une quantité : la régression

1. La régression linéaire
2. Mesurer l'erreur d'une régression
3. Régression polynomiale et régularisation
4. Atelier de synthèse — Prédire une quantité : la régression

06 Prédire une catégorie : la classification

1. Régression logistique et k plus proches voisins
2. Matrice de confusion, précision et rappel
3. Seuil de décision et classes déséquilibrées
4. Atelier de synthèse — Prédire une catégorie : la classification

07 Arbres de décision et méthodes d'ensemble

1. Arbres de décision
2. Forêts aléatoires
3. Boosting de gradient et comparaison des modèles
4. Atelier de synthèse — Arbres de décision et méthodes d'ensemble

08 Évaluer et optimiser un modèle

1. La validation croisée
2. Régler les hyperparamètres
3. Diagnostiquer un modèle
4. Atelier de synthèse — Évaluer et optimiser un modèle

09 Découvrir des structures : l'apprentissage non supervisé

1. Regrouper avec les k-moyennes
2. Réduire les dimensions avec l'ACP
3. Repérer des anomalies
4. Atelier de synthèse — Découvrir des structures : l'apprentissage non supervisé

10 Du modèle à l'usage : biais, limites et mise à disposition

1. Sauvegarder et réutiliser un modèle
2. Biais, limites et dérive des données
3. Documenter et présenter un modèle
4. Atelier de synthèse — Du modèle à l'usage : biais, limites et mise à disposition