

PRÉSENTATION DU PROGRAMME DE LA FORMATION

PARCOURS EXPERT EN INTELLIGENCE ARTIFICIELLE NIV.7 - 2 ANS YNOV

OBJET DE LA FORMATION

Formation préparant au Titre de "Ingénieur en science des données spécialisé en infrastructure data ou en apprentissage automatique" délivré par YNOV - RNCP39586 - NSF 326 - Niveau 7 enregistré au RNCP par décision de France Compétences en date du 01-10-2024

PRÉREQUIS POUR SUIVRE LA FORMATION ET OBTENIR LES QUALIFICATIONS AUXQUELLES ELLE PRÉPARE

Être titulaire d'un titre ou diplôme de niveau 4 et avoir suivi un parcours de formation de 5 ans visant un titre de niveau 7 dans le domaine de l'informatique

OU

Être titulaire d'un titre ou diplôme de niveau 5 dans le domaine de l'informatique et avoir suivi un parcours de formation de 3 ans visant un titre de niveau 7 dans le domaine de l'informatique dans le domaine de l'informatique

OU

Être titulaire d'un titre ou diplôme de niveau 6 dans le domaine de l'informatique et avoir suivi un parcours de formation de 2 ans visant un titre de niveau 7 dans le domaine de l'informatique

OBJECTIFS

Objectifs et contexte de la certification :

Le secteur stratégique de la data et de l'intelligence artificielle continue sa croissance. Les quantités de données numériques produites et traitées dans le monde sont exponentielles. L'enjeu est crucial pour les entreprises. Il s'agit de les valoriser et les exploiter pour produire des connaissances, faire des prédictions, analyser des comportements etc.

Cette certification professionnelle s'inscrit dans ce contexte et permet d'acquérir un socle de compétences communes nécessaires pour collecter, stocker, traiter et valoriser des données. Elle vise aussi à fournir les compétences nécessaires pour piloter des projets complexes, intégrer les problématiques des entreprises et élaborer des solutions adaptées qui permettront d'atteindre les objectifs fixés. Ces compétences sont complétées de compétences liées à l'apprentissage automatique ou à l'infrastructure data :

- Véritable expert en apprentissage automatique, il conçoit des modèles de Machine Learning et des indicateurs pour aider la prise de décision dans la stratégie d'entreprise. Il utilise ces modèles prédictifs pour faire des recommandations et aider les dirigeants dans leur pilotage d'entreprise.

- ou expert en architecture des données, il conçoit et administre les systèmes de collecte de données. Il assure la récupération, la mise à disposition et la sécurisation des données afin de permettre leur exploitation.

Activités visées :

Collecte de données structurées et non structurées - Stockage des données structurées et non structurées - Structuration, transformation et enrichissement des données - Sécurisation des données - Analyse des données - Visualisation des données, interprétation et communication des résultats - Support et accompagnement des utilisateurs - Élaboration et cadrage du projet DATA - Pilotage du projet DATA - Management d'équipe - Veille, éthique et gouvernance des données

Pour la spécialisation en infrastructure data s'ajoutent : Analyse des besoins et définition de l'architecture - Conception et déploiement de l'infrastructure DATA - Supervision et exploitation de l'infrastructure DATA - Maintenance et sécurisation de l'infrastructure DATA

Pour la spécialisation en apprentissage automatique s'ajoutent : Analyse du besoin et résolution de problèmes complexes - Développement de modèles d'apprentissage automatique - Déploiement et automatisation des modèles d'apprentissage automatique

Compétences attestées :

Élaborer une stratégie de collecte de données en définissant les données utiles et nécessaires pour répondre à une problématique, et en identifiant les sources des données afin de cadrer le travail à réaliser pour collecter les données ciblées.

Mettre en œuvre des techniques de collecte de données en exploitant les API externes et les bases de données disponibles, des techniques de web crawling et de web scraping afin de recueillir les données ciblées.

Automatiser la collecte de données en mettant en place des tâches planifiées et/ou des flux temps réel, en utilisant des logiciels d'automatisation afin de garantir l'actualisation des données.

Élaborer la stratégie de stockage des données avec un modèle de données adéquat en intégrant les différents types de données, l'utilisation envisagée (analyse, stockage, disponibilité, accessibilité) et le volume à stocker afin d'organiser le stockage des données.

Construire une base de données en sélectionnant la technologie (SQL, NoSQL), un système de gestion de base de données (SGBD) et/ou une solution de stockage BIG DATA, en assurant le paramétrage et l'implémentation afin de mettre en œuvre le modèle de données qui garantit la disponibilité et l'intégrité des données.

Sélectionner les technologies et les outils de traitement de données en identifiant les solutions existantes et en comparant leurs avantages et leurs inconvénients afin de traiter efficacement les données collectées.

Transformer les données à l'aide de langage de programmation ou en utilisant des outils dédiés (Talend, Spark...) afin d'obtenir des données nettoyées exploitables.

Développer un processus ETL en identifiant les bénéfices des technologies ETL (ex : facilité de développement), en exploitant la technologie ETL préalablement sélectionnée afin d'automatiser l'extraction, la transformation et le chargement de données.

Définir la politique de sécurisation des données en évaluant les risques, en qualifiant leur niveau de sensibilité, en identifiant les droits d'accès selon les rôles des différentes parties prenantes et en respectant les exigences légales (ex : RGPD) afin de garantir la bonne utilisation et l'intégrité des données.

Concevoir une architecture sécurisée et robuste, en intégrant des mesures de sécurité multicouches et des contrôles d'accès stricts, en mettant en place des solutions de protection des données (ex : chiffrement) pour sécuriser les données en transit et au repos, en veillant à l'anonymisation des données personnelles afin de répondre aux normes de protection de la vie privée et de sécurité des données.

Analyser les besoins métier et les enjeux exprimés par un commanditaire en réalisant des entretiens exploratoires et en récupérant les informations stratégiques nécessaires afin de cadrer le travail d'analyse des données à produire.

Définir les axes d'analyse et les métriques en identifiant les données à exploiter, celles disponibles et pertinentes pour traduire la problématique d'entreprise énoncée en problème numérique.

Réaliser des requêtes et des calculs en utilisant des outils de dashboarding, des tableurs, des requêtes SQL ou scripts Python afin de produire une analyse des données préalablement collectées.

Élaborer des modèles statistiques et des tests d'hypothèses en modélisant des relations entre les variables, en évaluant la pertinence des résultats des simulations afin de valider ou réfuter des hypothèses.

Représenter les données en choisissant les modèles de représentation les plus adaptés (ex :

histogramme, Heat map, nuage de points) et en utilisant des outils de représentation adaptés (ex : Office, power BI) afin de permettre la compréhension et l'exploitation des données par le public visé.

Présenter des recommandations, en préparant son discours et des arguments, en structurant son analyse sur les données représentées afin d'aider les décideurs à établir leurs stratégies.

Former les utilisateurs à l'utilisation des données et des outils de visualisation en analysant le besoin de montée en compétences et en élaborant des supports de formation et de sensibilisation adaptés afin de permettre aux utilisateur de maîtriser l'exploitation des données.

Rédiger la documentation technique d'utilisation du système d'analyse de données en identifiant le public concerné, en détaillant le fonctionnement du système d'analyse de données afin d'assurer la traçabilité et la transmission aux utilisateurs.

Définir les objectifs à atteindre et le périmètre du projet, en analysant les contraintes techniques et réglementaires, en étudiant le contexte et les enjeux afin de dimensionner le projet en termes de délai et budget.

Dimensionner le projet en évaluant la charge de travail et les ressources nécessaires (humaines, matérielles) au regard des exigences attendues et des contraintes préalablement définies afin d'estimer le temps et le budget nécessaires à la faisabilité du projet.

Rédiger la documentation projet, en identifiant les parties prenantes concernées, en prenant en compte l'ensemble des caractéristiques du projet, afin de clarifier et formaliser les attendus.

Planifier l'exécution du projet en organisant la répartition et l'ordonnancement des activités, le planning prévisionnel de réalisation et les ressources nécessaires à son exécution, en prenant en considération les personnes en situation de handicap afin de suivre les différentes phases du projet.

Suivre l'avancement du projet en mettant en place un outil de suivi (logiciel de suivi, tableau de bord), en définissant les indicateurs (qualitatifs et/ou quantitatifs) pour chaque jalon défini dans le planning, en réalisant des reportings et des comptes rendus de réunion afin d'anticiper les aléas éventuels.

Évaluer les besoins en compétences de l'équipe projet, en collaborant avec le service Ressources Humaines, en établissant un plan de développement des compétences et en orientant les membres de l'équipe vers des formations adaptées, afin de renforcer l'équipe responsable de mener à bien le projet DATA.

Piloter l'équipe projet en affectant les missions à réaliser, en prenant en compte les spécificités des membres de l'équipe, en intégrant les spécificités d'un contexte multiculturel, international, en utilisant les différentes techniques de communication et d'animation managériale pour favoriser le bon fonctionnement de l'équipe.

Procéder aux arbitrages et aux réajustements nécessaires à partir de l'analyse des écarts entre le prévisionnel et l'état du projet à date, en utilisant des outils d'aide à la décision (ex : logigramme) afin de garantir la performance du projet dans le respect des objectifs de qualité, coûts et délai.

Mettre en place un système de veille technologique et réglementaire en matière de science des données et d'Intelligence Artificielle à l'aide de recherches documentaires, de plateformes de partage, de webinars afin d'être alerté des évolutions qui impacteraient les pratiques métier.

Intégrer dans ses pratiques métier les enjeux en termes de données responsables, de responsabilité sociétale et environnementale (RSE), de sécurité, d'éthique et de confidentialité des données en se tenant informé des évolutions du cadre juridique, à travers une recherche documentaire ou en étant accompagné par des juristes afin d'agir dans le respect de la législation.

Pour la spécialisation en infrastructure data :

Analyser l'environnement du projet en recueillant les besoins métiers, les volumes de données à traiter, en réalisant un état des lieux des composants existants afin d'orienter le choix de conception de l'architecture DATA à mettre en œuvre.

Sélectionner l'ensemble des composants et technologies de l'infrastructure en étudiant les solutions existantes, en vérifiant leur compatibilité et les normes en vigueur afin de concevoir une architecture DATA correctement dimensionnée pour le projet.

Concevoir une architecture d'entrepôt de données en s'appuyant sur le cahier des charges, en sélectionnant les composants appropriés afin d'optimiser le stockage des données en termes de rapidité, de sécurité et d'accessibilité.

Mettre en place des pipelines DATA temps réel ou asynchrones à l'aide d'outils BIG DATA (ex : DATA brick, plateforme DBT Snowflake), Cloud et/ou On premise afin d'automatiser la transformation et la transmission les données.

Automatiser l'intégration et le déploiement des composants en utilisant des outils d'intégration et de développement continu, afin d'industrialiser la mise en production de l'architecture DATA.

Mettre en place un système de supervision et d'alertes en déployant des outils de supervision, et en déterminant les indicateurs de suivi pertinents afin de s'assurer du bon fonctionnement des composants de l'architecture DATA.

Exploiter les systèmes et les équipements de l'infrastructure DATA en respectant les procédures d'administration et de maintien en condition opérationnelle afin de garantir l'intégrité et la disponibilité des données de l'organisation.

Rédiger la documentation fonctionnelle et technique en identifiant le public concerné et les objectifs attendus, en prenant en compte l'ensemble des caractéristiques de l'infrastructure DATA afin de transférer les procédures et les modalités de fonctionnement aux différents utilisateurs.

Élaborer le cahier de recette en rédigeant les scénarios de tests et les résultats attendus afin de détecter les anomalies de fonctionnement et les régressions éventuelles.

Résoudre un incident technique en investiguant la source du problème et en déployant une méthode de résolution afin de rétablir la disponibilité du service.

Pour la spécialisation en apprentissage automatique :

Analyser la problématique et le contexte d'un commanditaire en réalisant des entretiens exploratoires, des questionnaires et une analyse de l'existant afin de lui apporter une réponse appropriée.

Cadrer la stratégie de résolution du problème, en utilisant des algorithmes, en traduisant le problème en un problème d'optimisation afin de le résoudre avec les outils des modèles d'apprentissage automatique.

Sélectionner les technologies, les outils et les algorithmes en identifiant les différentes solutions disponibles et en comparant leurs avantages et leurs inconvénients afin de répondre à la problématique du commanditaire au regard des contraintes du projet.

Construire des variables en utilisant des langages de programmation (ex : Python, Scala, R, Julia...) en exploitant des bibliothèques d'analyse de données afin de fournir les meilleures variables au modèle d'apprentissage automatique.

Sélectionner les variables en identifiant les différentes méthodes de sélection de variables possibles, en utilisant des méthodologies d'apprentissage automatique afin d'optimiser la performance du modèle.

Entrainer un modèle d'apprentissage automatique à l'aide de librairies (ex : Scikit-learn, XGBoost, TensorFlow, PyTorch) afin d'obtenir des modèles capables de prédictions sur de nouvelles données inconnues.

Optimiser la performance des modèles d'apprentissage automatique en modifiant les hyperparamètres et en analysant les prédictions afin de répondre au mieux à la problématique du commanditaire.

Sauvegarder le modèle d'apprentissage automatique entraîné à l'aide d'outils de sérialisation, virtualisation, containerisation, versioning afin de pouvoir le déployer dans des environnements de production.

Déployer des modèles d'apprentissage automatique en utilisant des API et des outils CI/CD afin de le mettre en production.

Superviser le système Machine Learning en sélectionnant des outils de monitoring et en les exploitant afin de détecter les dérives et les bugs du modèle d'apprentissage automatique.

Automatiser les tâches inhérentes au cycle de vie d'un système d'apprentissage automatique à l'aide de pipelines et des outils adaptés afin de maintenir la performance du modèle d'apprentissage automatique.

DURÉE ET CONTENU DE LA FORMATION

Le volume horaire total de la formation s'élève à 952 heures. Il est découpé comme suit :

Programme de première année

Le programme comprend 476 heures

Les 476 heures sont réparties en 6 Unités Pédagogiques, dont 40 heures de FOAD.

UNITÉS PÉDAGOGIQUES Unités d'enseignements	SÉQUENCES	TOTAL HEURES
Se challenger et développer ses compétences	Concours Natio IA & Data	93
	Hackaton IA IA & Data	
	Epreuves formatives [2 - SEMESTRE 2]	
	Epreuves formatives [2 - SEMESTRE 2]	
	Ydays	
Objectif Pro	Objectif Pro - Approche par compétences	15 dont 1 FOAD
	Objectif Pro - Identité Pro	
	Objectif Pro - Posture	
	Objectif Pro - Soft Skills	
Comprendre et interagir avec son environnement	Intelligence Artificielle	51 dont 39 FOAD
	Welcome Day	
	Impact & RSE	
	Anglais	
MAJEURES	Data science	189
	Machine Learning	
	Algorithmie et développement dans l'ingénierie des données	
MINEURES	Gestion d'un projet IT	84
	Outils ETL	
	Pilotage de la stratégie SI	
Tronc commun	Découverte des épreuves certificatives	44
	MasterClass - 4h avec votre directeur	
	hackathon M1	
	Culture des entreprises	
Total		476 dont 40 FOAD

Programme de seconde année

Le programme comprend 475 heures et 1 heures d'évaluation

Les 476 heures sont réparties en 6 Unités Pédagogiques, dont 40 heures de FOAD.

UNITÉS PÉDAGOGIQUES Unités d'enseignements	SÉQUENCES	TOTAL HEURES
Comprendre et interagir avec son environnement	Préparation & Epreuves certificatives	69 dont 39 FOAD
	Welcome Day	
	Anglais	
	Intelligence Artificielle	
	Impact & RSE	
MINEURES	MLOps	84
	Data visualisation M2	
	Large Language Model	
Se challenger et développer ses compétences	Ydays	84
	Concours Natio Info	
Objectif Pro	Objectif Pro - Approche par compétences	18 dont 1 FOAD
	Objectif Pro - Posture	
	Objectif Pro - Soft Skills	
	Objectif Pro - Identité Pro	
	Objectif Pro - Après Ynov	
M2 - TC Filière IT	Préparation aux épreuves certificatives	32
MAJEURES	Industrialisation de l'IA dans le cloud	189
	Deep Learning	
	Natural Language processing	
Total		476 dont 40 FOAD

MODALITÉS ET DÉLAIS D'ACCÈS

Le parcours de formation est couvert par un Contrat d'apprentissage ou un Contrat de professionnalisation dont la date du début d'exécution est le : 29/09/2025

MÉTHODES MOBILISÉES ET MODALITÉS D'ÉVALUATION

Modalités d'évaluations formatives :

Afin de permettre d'évaluer les acquis de la formation par chaque apprenant, des sessions d'évaluations sont réparties de la manière suivante :

- Un contrôle continu au sein de chaque module selon les modalités les plus adaptées aux objectifs poursuivis : QCM, exercices, étude de cas, ateliers de mise en pratique, mises en situations, techniques d'observation etc. ;
- D'éventuelles présentations de projets transverses à l'écrit ou à l'oral, réalisées de manière individuelle ou collective ;
- Un accompagnement pédagogique lors de la période en entreprise. Les résultats des évaluations sont mis à disposition des apprenants au fur et à mesure des corrections.

Les résultats des évaluations sont mis à disposition des apprenants au fur et à mesure des corrections.

Modalités d'évaluations certificatives :

L'évaluation de la certification est réalisée par bloc de compétences, les modalités des évaluations certificatives étant décrites dans les règlements de certification du Titre préparé.

MOYENS PÉDAGOGIQUES

Méthodes pédagogiques mobilisées :

La formation mobilise plusieurs méthodes pédagogiques afin de permettre aux bénéficiaires d'acquérir l'ensemble des compétences sanctionnées par le titre visé.

Parmi les modules de compétences et les cours, un scenario pédagogique générique, pensé uniquement en fonction des besoins et des contraintes du public cible, des objectifs pédagogiques et des ressources humaines disponibles, a été conçu. Ce scénario sera adapté en début de formation pour tenir compte des contraintes inhérentes à l'activité. Les méthodes pédagogiques mobilisées :

- Méthode dite « expositive » au travers de l'utilisation des présentations permettant l'assimilation des connaissances requises pour l'exercice de la fonction.
- Méthode démonstrative au travers de mise en situation professionnelle et de jeux de rôles permettant l'acquisition de compétences par la « reproduction » régulière d'actions,
- Méthode active : au travers d'un mémoire, de divers projets de groupe, de réponse à des problématiques réelles d'entreprise (en mode challenge) le bénéficiaire

s'engage dans l'acquisition de connaissances et construit son socle de compétences professionnelles.

La FOAD est déployée sur certains cours, elle s'inscrit et est précisée dans le syllabus des cours en question, elle se décline dans le plan de progression pédagogique du cours. Les modalités de déroulement, de suivi et de sanction de l'action ainsi que les critères d'évaluation sont communiqués au formateur au travers du syllabus ainsi qu'aux bénéficiaires de la formation. Il contribue à la réalisation et au rendu de travaux donnant lieu à une évaluation.

Dans le cadre de l'utilisation de LMS, l'assistance technique est réalisée par le service pédagogique de l'école, ainsi chaque alternant peut notamment se connecter via son extranet et bénéficier de tous les documents et supports pédagogiques mis à disposition. L'assistance pédagogique est assurée par le formateur en charge du cours/projet.

Les activités pédagogiques mobilisées dans le cadre de la FOAD sont :

- Des outils e-learning,
- Des ateliers en autonomie,
- Des challenges et autres séminaires,
- Des projets