

Diplôme d'Université



Analyste Data Science

Le phénomène « Big Data » et les nouvelles données issues du web, des réseaux sociaux ou encore des objets connectés constituent un défi majeur en informatique décisionnelle et en statistique pour les prochaines années. Il s'agit de considérer les données accumulées dans les systèmes d'information ou disponibles via des objets connectés comme un capital qu'il faut chercher à valoriser, à visualiser et piloter, en leur appliquant différents traitements informatiques et statistiques. Les défis à surmonter pour atteindre ces objectifs passent par la prise en compte de données volumineuses, arrivant souvent sous forme de flux continus et de provenance ou de type hétérogène, ce qui implique diverses facettes de la « Data Science » et le développement de nouveaux outils répondant à ces défis, aussi bien au niveau des données que dans leur traitement statistique dans une perspective décisionnelle.

Quel type de formation ?

Formation continue

Comment s'organisent les études ?

► Formation continue

Le DU est constitué de 5 certificats (modules) qui alternent enseignements théoriques, travaux dirigés et travaux pratiques.

Le DU comprend 12 semaines de formation. Il se déroule avec un rythme de 2 jours de cours toutes les 2 semaines, afin de faciliter le suivi par des professionnels en entreprise. Chaque certificat se déroule ainsi sur 4 semaines.

Les modules s'appuieront sur des outils support et des partenariats avec des professionnels du secteur. Dans chaque module seront aussi étudiées des applications d'un ou plusieurs domaines parmi lesquels on peut citer : les télécommunications, la finance, l'énergie, les réseaux sociaux, la relation-client et le marketing.

Objectifs

Ce Diplôme Universitaire propose un complément de formation aux nouveaux concepts et outils pour le « Machine Learning » et la « Data Science », pour des professionnels ayant une formation de base en statistique et/ou en informatique décisionnelle (bases de données, statistique, fouille de données).

Compétences visées

- Connaître la syntaxe des langages statistiques et savoir l'utiliser.
- Analyser, structurer et visualiser statistiquement des données tant simples que complexes.
- Comprendre les spécificités des données complexes et de leur exploitation, y compris en temps réel.
- Mettre en œuvre une analyse dans le cas de données

spécifiques (temps réel, spatio-temporelles, textuelles, graphes et réseaux).

- Évaluer et améliorer les performances d'un modèle prédictif.

Moyens pédagogiques

Alternance d'enseignements théoriques, travaux dirigés et travaux pratiques sur des outils du marché.

Perspectives professionnelles

Ce diplôme permettra aux apprenants d'évoluer vers des postes au sein de projets « Data Science » ou « Big Data » dans les entreprises, les administrations et les collectivités territoriales. Il permettra également d'accompagner au niveau technique ces entités dans les évolutions liées à la révolution digitale.

Comment nous rejoindre ?

Conditions d'admission

Salariés ou professionnels souhaitant valider et compléter des acquis professionnels dans le domaine du traitement de l'information, ayant un niveau Bac +2 avec des compétences en statistique et informatique décisionnelle.

Procédure d'admission

- Candidature en ligne :

Session de candidature 1 :

27/05/2024 au 01/07/2024

<https://ecandidat.app.u-paris.fr/sh1/>

Session de candidature 2 :

26/08/2024 au 02/12/2024

Plus d'infos sur le site de l'IUT

- Sélection sur dossier et entretien par l'équipe pédagogique.
- Coût du cycle de formation continue : 3 000,00 €
- La prise en charge du coût de la formation peut être assurée dans le cadre des dispositifs de Formation Professionnelle en vigueur. Il appartient aux candidats d'effectuer les démarches nécessaires auprès des organismes dont ils dépendent : DRH, OPCO, Pôle Emploi...

Droits d'inscription universitaire :

Obligatoires, ils sont fixés chaque année par arrêté à paraître en juillet.

Le contenu pédagogique

La pratique du contrôle continu des connaissances pour l'attribution du diplôme implique la participation des stagiaires à tous les cours.

MODULE 1 - APPRENTISSAGE STATISTIQUE

Data Mining avec R

Arbres, Forêts aléatoires, Boosting

- Explorer le Data mining avec R : les principales méthodes classiques de la fouille de données (clustering, classification, prévision) avec leur pré-traitements ainsi qu'un langage informatique adapté
- Se former aux méthodes modernes d'apprentissage statistique supervisé (dont les méthodes d'arbres, les forêts aléatoires, les méthodes d'ensembles et le boosting)

- Pré-traitements des données complexes (transformations, indexation, signatures,...)

- Méthodes de fouille de séries temporelles, de séries fonctionnelles, de données spatiales, de textes

MODULE 4 - FOUILLE DU WEB

ET DES RESEAUX SOCIAUX

Extraction, Visualisation de données du web et des réseaux

Analyse et propriétés statistiques des grands réseaux

- Sources et formats de données du web
- Extraction de données à partir du web
- Structures de grands graphes, propriétés statistiques de ces graphes
- Algorithmes de fouille des graphes, des réseaux sociaux (statique, dynamique)
- Visualisation de graphes

MODULE 2 - DÉCISIONNEL ET VISUALISATION EN TEMPS RÉEL

Flux de données et méthodes de fouille associées

Objets connectés, données capteurs et visualisations dynamiques

- Sources de données en flux, données capteurs et objets connectés.
- Tableaux de bord dynamiques, alarmes, visualisation de flux de données
- Méthodes de fouille de flux de données : modèles incrémentaux et adaptatifs (clustering, classification, prévision)

MODULE 5 - BIG DATA ET IA

Interfaçage des données massives avec Python

Réseaux de neurones et deep learning

- Bases de données massives, structuration et interfaçage avec les outils logiciels du data mining et de l'apprentissage statistique ainsi qu'un langage informatique adapté (Python)
- Les nouveaux outils d'apprentissage : les réseaux de neurones et le deep learning

MODULE 3 - FOUILLE DE DONNÉES COMPLEXES

Text Mining

Analyse et prévision de séries temporelles

Fouille de données spatiales

- Sources de données complexes : séries temporelles, données spatiales, texte

Comment nous contacter ?

Obtenir un devis et remplir le dossier de financement :

Service Formation Continue et Alternance

- 01 76 53 49 75
- bernadette.amiaud@u-paris.fr

Obtenir les informations pédagogiques :

Département Science des Données :

- 01 76 53 49 78
- sophie.mauray@u-paris.fr

Les cours sont assurés sur le site Pajol

20 quater, rue du Département
75018 Paris

- M La Chapelle, Stalingrad, Gare du Nord, Marx Dormoy
- RER Ligne B et D station Gare du Nord
- T T3 station Porte de la Chapelle

IUT de Paris - Rives de Seine

143 avenue de Versailles

75016 Paris

Tél : 01 76 53 47 00

iutparis-seine.u-paris.fr

