

Certificat
Data Science pour la Détection de la Fraude Financière
Data Science for Financial Fraud Detection

1. Objectifs

Un certificat est une formation diplômante courte, centrée sur une problématique métier, et qui donne lieu à la délivrance d'un Diplôme d'Université (DU), à l'issue d'une évaluation.

Le certificat « Data Science pour la Détection de la Fraude Financière » a pour objectif de former les participants aux méthodes économétriques et d'apprentissage automatique supervisée et non-supervisée dans le contexte spécifique de la détection de la fraude financière et la lutte contre le blanchiment d'argent. Ces techniques peuvent être appliquées pour la détection d'une grande variété de fraudes, telles que la fraude à la carte, la fraude à l'assurance, la fraude dans le secteur de la santé, la fraude fiscale, le blanchiment d'argent, etc. Le cours présente à la fois les fondements théoriques des méthodes ainsi que des illustrations pratiques sous différents logiciels statistiques.

La formation sera complétée par la réalisation et la présentation d'un projet de mise en œuvre des techniques acquises pendant le cours en lien avec les problématiques de l'entreprise dans laquelle travaillent les participants.

2. Programme

Le programme du certificat sera structuré autour de 5 modules :

Module 1 : Détection de la fraude - éclairages juridiques (2 jours)

- Fondamentaux
 - Définition, typologies des fraudes, victimes et auteurs
 - Profil du fraudeur
- Enjeux et risques liés à la fraude & Sensibilisation du public
 - Enjeux sociaux, économiques
 - Image, réputation, respect des procédures
- Architecture institutionnelle
 - Cadre national, européen, international
 - Lutte contre la fraude : réglementations existantes et à venir
 - Lutte contre le blanchiment d'argent : réglementations existantes et à venir
 - Fraude et blanchiment d'argent : une lutte commune
- GDPR
- Prévention, détection et cartographie des risques
- Contrôles et sanctions
- Retour d'expériences
 - Cas récents de fraude aux États-Unis et en Europe et enseignements tirés

Module 2 (2 jours)

Module 2.1 : Gestion et traitement des données (1 jour)

- Data
 - Sources des données
 - Prétraitement des données
 - Types de variables
 - Détection et traitement des valeurs manquantes
 - Détection et traitement des valeurs extrêmes
 - Transformation des variables
 - Sélection des variables, etc.

- Enjeux spécifiques de la modélisation de la fraude
 - Endogénéité
 - Asymétrie
 - Instabilité, etc.

Module 2.2 : Méthodes non-supervisées pour la détection de la fraude (1 jour)

- Introduction : principe et définition
- Synthèse sur les approches non-supervisées de détection de la fraude
 - Présentation de principales méthodes non-supervisées
 - Méthodes graphiques et statistiques
 - Méthodes de classification
 - Loi de Benford
 - Avantages et limites
- Étude de cas

Module 3 : Méthodes supervisées pour la détection de la fraude (2 jours)

- Introduction :
 - Principe et définition
 - Économétrie versus Machine Learning
- Synthèse sur les approches supervisées de détection de la fraude
 - Méthodes économétriques :
 - Régressions linéaires et logistiques
 - Avantages et limites
 - Méthodes d'apprentissage automatique (Machine Learning) :
 - Méthodes classiques : arbres de décision, réseaux de neurones, SVM
 - Méthodes combinées : Bagging, Boosting et Random Forests
 - Extensions : Boosted Decision Trees, XGBoost et LinXGBoost
 - Avantages et limites
- Étude de cas

Module 4 (2 jours)

Module 4.1 : Évaluation des modèles de détection de la fraude (1 jour)

- Mesures de performance
 - AUC
 - LIFT, etc.
- La fraude comme « événement rare »
 - Présentation des méthodes de ré-échantillonnage
- Compréhension des « boîtes noires » derrière les méthodes d'apprentissage automatique
 - Introduction à la méthode LIME
- Étude de cas

Module 4.1 : Les réseaux sociaux et la fraude (1 jour)

- Présentation des réseaux sociaux et les mécanismes de fraude
- Gestion des alertes
- Étude de cas

Module 5 : Présentation d'un projet (1 jour)

Le dernier jour de formation sera consacré (i) à la présentation d'un projet de détection de la fraude par les participants sur une problématique inhérente à leur entreprise et (ii) à une conférence sur le thème de la détection de la fraude.

Ce programme est donné à titre indicatif et est susceptible de subir des modifications mineures.

3. Intervenants

Responsable de la formation

Denisa Banulescu-Radu est maître de conférences à l'université d'Orléans et membre de l'équipe économétrie du Laboratoire d'Économie d'Orléans (LEO, FRE CNRS 2014). Spécialisée en économétrie financière, ses recherches portent principalement sur la détection de la fraude financière et la lutte contre le blanchiment d'argent, la modélisation de la volatilité financière et la validation des modèles de risque. Denisa est par ailleurs membre du Cercle de la Compliance et lauréate du prix Jeune Chercheur de l'Autorité des Marchés Financiers (2016) et du prix de Thèse de la Fondation Banque de France (2015).

Sandie Lacroix-De Sousa est maître de conférences à l'université d'Orléans et membre du Centre de Recherche Juridique Pothier (CRJP, EA 1212). Spécialisée en droit des affaires, ses recherches portent principalement sur les obligations et responsabilités du banquier, la lutte contre le blanchiment d'argent et le droit de l'ingénierie sociétaire. Elle est par ailleurs codirectrice du Master 2 Droit et ingénierie du patrimoine et enseigne le droit bancaire et financier à l'Université Jagellonne de Cracovie. Membre du Cercle de la Compliance, sa thèse a reçu le prix de la Fondation Varennes dans la catégorie droit des activités économiques (2011).

Intervenants

La formation est assurée par des professionnels et des universitaires spécialisés dans la détection de la fraude, les méthodes de machine learning et la réglementation.

A COMPLETER

4. Admission

Public

Ce certificat s'adresse à un public de professionnels qui souhaitent acquérir une spécialisation dans les méthodes économétriques et de machine learning spécifiquement adaptées au contexte de détection de la fraude financière. Le public ciblé est constitué de professionnels (Data Scientists, ingénieurs statisticiens, chargés d'études statistiques, consultants, compliance officers, etc.) travaillant au sein de directions des risques de banques ou de sociétés d'assurances, ou au sein de cabinets de conseil spécialisés dans la détection de la fraude et la lutte contre le blanchiment d'argent.

Des connaissances en programmation de base sous Python et R sont recommandées pour suivre les applications de ce certificat.

Candidature

Pour candidater à la formation, veuillez envoyer un email à XXX@univ-orleans.fr avec un CV ou un profil LinkedIn joint. Le nombre de participants sera limité à 10 par session.

Les compétences acquises en entreprise, la qualité des études et/ou l'adéquation de l'expérience professionnelle au programme du certificat constituent les principaux critères de sélection.

Pour plus d'informations, consulter le site [XXXXXXX](#) .

5. Informations pratiques

Organisation

L'organisation des cours permet la continuité de votre activité professionnelle. La durée de la formation est de 9 jours.

La formation se déroulera sur 4 mois et comportera 8 jours de formation (6 heures par jour) répartis en 4 sessions de 2 jours consécutifs, à raison d'une session toutes les 3 à 4 semaines. Le dernier jour de formation est consacré à la présentation des projets réalisés par les participants et à une conférence.

Cette formation est dispensée en français. Les supports utilisés seront rédigés en français et/ou en anglais.

Agenda de la formation

La session d'automne du certificat aura lieu aux dates suivantes

- Module 1 : Jeudi XX et vendredi XX
- Module 2 : Jeudi XX et vendredi XX
- Module 3 : Jeudi XX et vendredi XX
- Module 4 : Jeudi XX et vendredi XX
- Module 5 : Vendredi XX

Tarif de la formation

Le tarif de la formation est de 4800€ net de taxe. L'Université d'Orléans n'est pas assujettie à la TVA pour ses actions de formation. Ce tarif comprend la mise à disposition du matériel pédagogique et la prise en charge des pauses repas du midi.

Un package incluant le prix de la formation, les transports (5 trajets SNCF aller-retour, 2nde classe, Paris Gare d'Austerlitz - Gare d'Orléans Centre), l'hébergement pour 4 nuits (une nuit par session de formation de deux jours) est proposé pour un tarif de 5600€ net de taxe.

Lieu de la formation

La formation sera dispensée à l'[hôtel Dupanloup](#), situé dans le centre-ville d'Orléans à quelques minutes à pied de la gare SNCF d'Orléans.

Alliance du passé et du futur, ce joyau architectural du XVII^{ème} siècle fut jusqu'en 1905 la résidence des évêques d'Orléans. Totalement réhabilité en 2014, l'hôtel Dupanloup accueille aujourd'hui le Centre International Universitaire pour la Recherche de l'Université d'Orléans.

Transport (depuis Paris) et hébergement

Pour se rendre sur le lieu de la formation, comptez environ une heure et quart à partir de la gare d'Austerlitz. La durée du trajet en train de la gare d'Austerlitz à la gare d'Orléans Centre est d'environ une heure, à laquelle il convient d'ajouter environ 10 minutes de trajet à pied pour se rendre à l'hôtel Dupanloup. Pour les intervenants qui souhaitent rester sur Orléans le premier soir de chaque session de formation, nous communiquerons une liste d'hôtels proches de lieu de la formation. Les participants ayant opté pour le package incluant les transports et l'hébergement seront hébergés (dans la mesure des places disponibles) à l'[hôtel Abeille](#).

Intervenants

La formation est assurée par des professionnels et des universitaires spécialisés dans la détection de la fraude, les méthodes de machine learning et la réglementation.

Florence **Giuliano** – SAS

Sandie **Lacroix-De Sousa** – CRJP

A COMPLETER