

LA DÉTECTION DE LA FRAUDE FINANCIÈRE - UNE VISION ACADÉMIQUE -

**LE 1ER FORUM RÉGIONAL SUR
L'INTELLIGENCE ARTIFICIELLE AU SERVICE DES ENTREPRISES
LAB'O, LE 16 OCTOBRE 2019**

Denisa BANULESCU-RADU

MAÎTRE DE CONFÉRENCES – LEO, UNIVERSITÉ D'ORLÉANS



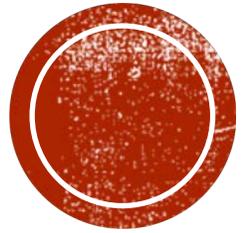


BACK
TO
SCHOOL

1. L'écosystème académique
2. Pourquoi la détection de la fraude ?
3. **RedFlag** : Appel à Projets de recherche d'initiative académique
4. Diplômes d'Université

SOMMAIRE





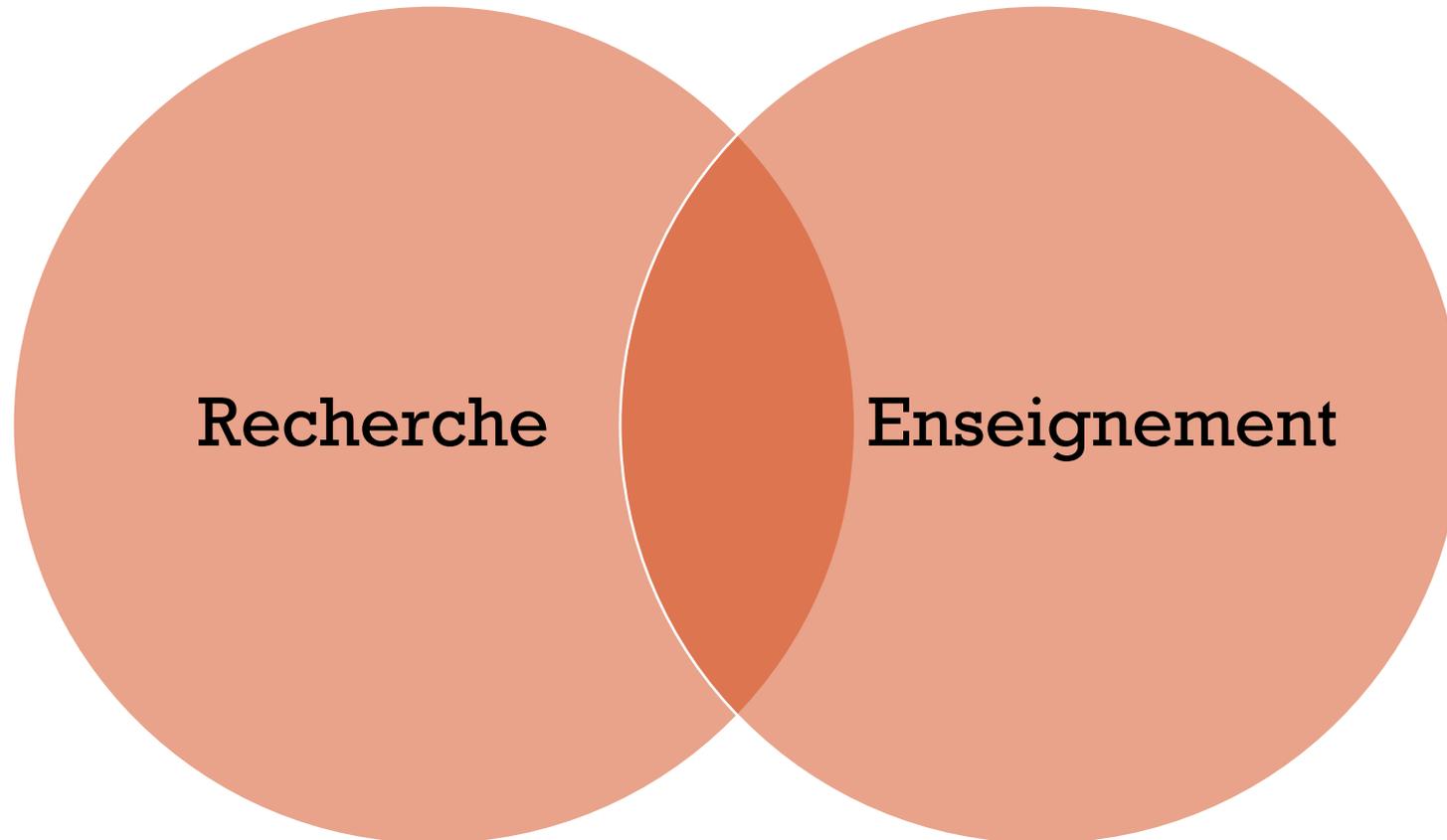
L'ÉCOSYSTÈME ACADÉMIQUE

L'ÉCOSYSTÈME ACADÉMIQUE



LÉO
Laboratoire
d'Économie
d'Orléans

Unité Mixte de Recherche
n° 7322 du CNRS



Université d'Orléans



LABORATOIRE D'ÉCONOMIE D'ORLÉANS



LÉO
Laboratoire
d'Économie
d'Orléans

Unité Mixte de Recherche
n° 7322 du CNRS

- Créé le 1^{er} Janvier 1996
- L'unique laboratoire de recherche en économie en région Centre-Val de Loire sous la tutelle conjointe de **l'Université d'Orléans**, de **l'Université de Tours** et du **CNRS**
- Une centaine de membres
- Trois domaines de compétences :
 - Macroéconomie et Finance
 - Économétrie
 - Économie Internationale et Développement Durable (EI2D)



212

articles de
recherche



40

conférences



28

contrats de
recherche



52

professeurs
invités



4 prix de thèse

1 prix AMF



MASTER ÉCONOMÉTRIE ET STATISTIQUE APPLIQUÉE



Université d'Orléans

- Créé en 2004
- 350 diplômés
- Environ 700h de formation (dont 564h sous SAS)
- Equipe pédagogique
 - 5 professeurs + 4 MCF
 - Professionnels

Exemple: **Fraud and Risk - Florence Giuliano** (Fraud strategy director, SAS EMEA)
- Plusieurs stages de fin d'études dans le domaine de la **détection de la fraude et la lutte contre le blanchiment d'argent**



MASTER ÉCONOMÉTRIE ET STATISTIQUE APPLIQUÉE

Semestre 3

Unité d'enseignement	COEF. / ECTS	CM (volume horaire)	TD (volume horaire)
Master 2 Semestre 9			
Statistique non paramétrique	2	12	
Méthodes de scoring	4	24	
Advanced Financial Econometrics	4	24	
Econométrie semi et non paramétrique	2	12	
Modèles de durée	4	24	
Panel Data Econometrics	2	12	
Big Data analytics: trees & aggregation methods (bagging, random Forests & boosting)	2	12	
Big Data analytics: penalized regressions (Lasso, Adaptive Lasso, Elastic-Net)	2	12	
Big Data analytics: Support Vector Machine	2	12	
Réglementation prudentielle bancaire	2	12	
Détection de la fraude	2	12	
Communication orale	2	9	
Cours du partenariat SAS: SAS IML, SAR OR, SAS IML PLUS...			

Semestre 4

Unité d'enseignement	COEF. / ECTS	CM	TD
Master 2 Semestre 10 option professionnelle			
Data Mining	4	24	
Assurance et techniques actuarielles 2	2	12	
Introduction au langage VBA	3	15	
Gestion de bases de données sous SAS	2	12	
Mise en œuvre de la Proc SQL sous SAS	2	12	
Méthodes de prévision	2	12	
Big Data analytics: neural networks	2	12	
Machine Learning with Python	3	12	
Évaluation des enseignements par les étudiants			1
Stage (3 mois minimum)	10		



MASTER ÉCONOMÉTRIE ET STATISTIQUE APPLIQUÉE

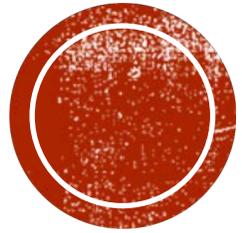
Semestre 3

Unité d'enseignement	COEF. / ECTS	CM (volume horaire)	TD (volume horaire)
Master 2 Semestre 9			
Statistique non paramétrique	2	12	
Méthodes de scoring	4	24	
Advanced Financial Econometrics	4	24	
Econométrie semi et non paramétrique	2	12	
Modèles de durée	4	24	
Panel Data Econometrics	2	12	
Big Data analytics: trees & aggregation methods (bagging, random Forests & boosting)	2	12	
Big Data analytics: penalized regressions (Lasso, Adaptive Lasso, Elastic-Net)	2	12	
Big Data analytics: Support Vector Machine	2	12	
Réglementation prudentielle bancaire	2	12	
Détection de la fraude	2	12	
Communication orale	2	9	
Cours du partenariat SAS: SAS IML, SAR OR, SAS IML PLUS...			

Semestre 4

Unité d'enseignement	COEF. / ECTS	CM	TD
Master 2 Semestre 10 option professionnelle			
Data Mining	4	24	
Assurance et techniques actuarielles 2	2	12	
Introduction au langage VBA	3	15	
Gestion de bases de données sous SAS	2	12	
Mise en œuvre de la Proc SQL sous SAS	2	12	
Méthodes de prévision	2	12	
Big Data analytics: neural networks	2	12	
Machine Learning with Python	3	12	
Évaluation des enseignements par les étudiants			1
Stage (3 mois minimum)	10		





POURQUOI LA DÉTECTION DE LA FRAUDE ?

DÉFINITION

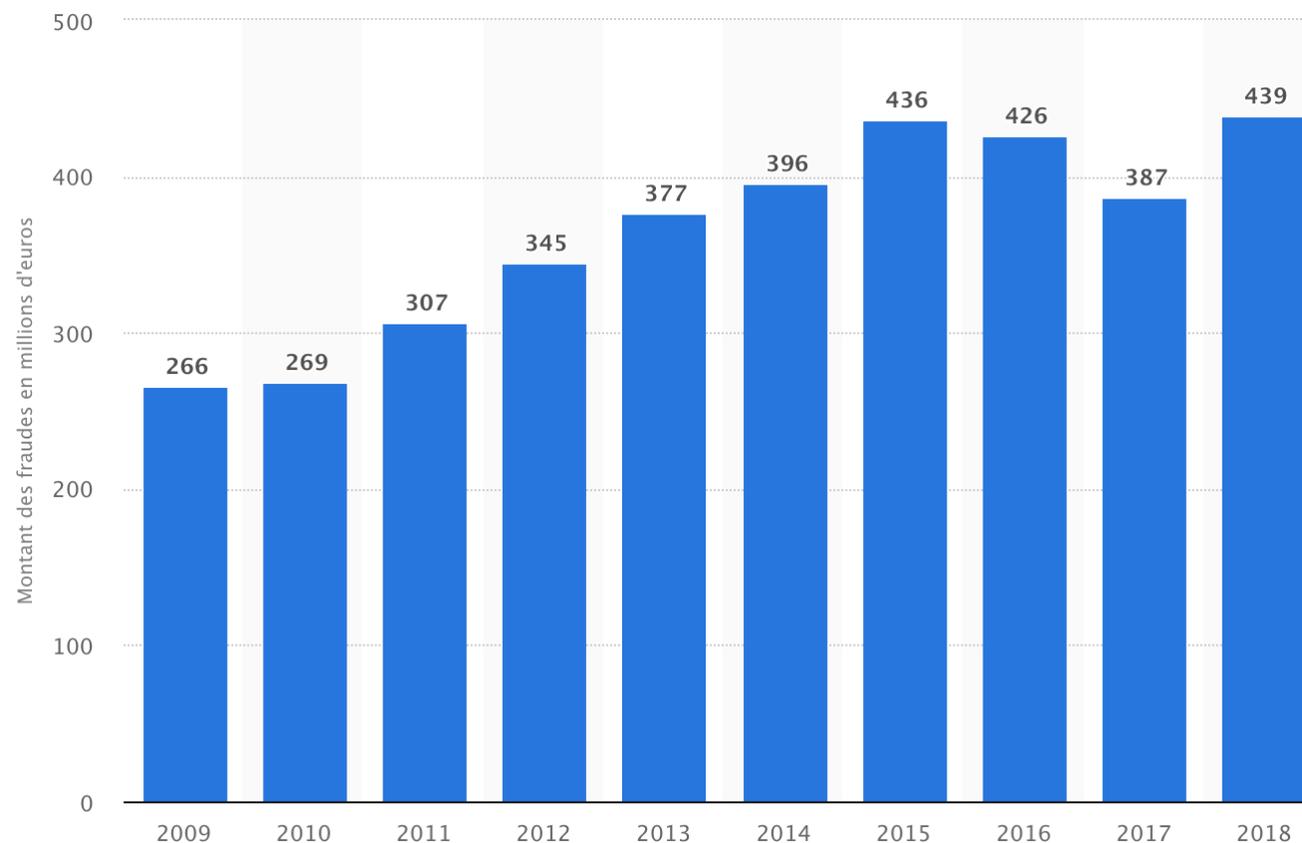
- *La FRAUDE est un crime rare, bien réfléchi, dissimulé de manière imperceptible, évoluant dans le temps et souvent soigneusement organisé, qui se présente sous de nombreuses formes.* (Baesens et al., 2015)



TYPES DE FRAUDE ET CHIFFRES CLÉS

▪ Fraude à la carte

439 M€ sur 569,5 Md€
soit environ 8 centimes fraude
pour 100€ *



Valeur des fraudes à la carte bancaire en France de
2009 à 2018 (en millions d'euros) *

* source : [statista.fr](https://www.statista.fr)



TYPES DE FRAUDE ET CHIFFRES CLÉS

- Fraude à la carte
- **Fraude à l'assurance**



Record: 2,5 Md€ en 2014,
dont 219 M€ récupérés *



TYPES DE FRAUDE ET CHIFFRES CLÉS

- Fraude à la carte
- Fraude à l'assurance
- **Fraude sociale et fiscale**



20 Md€ en 2018*

* source : tacotax.fr



TYPES DE FRAUDE ET CHIFFRES CLÉS

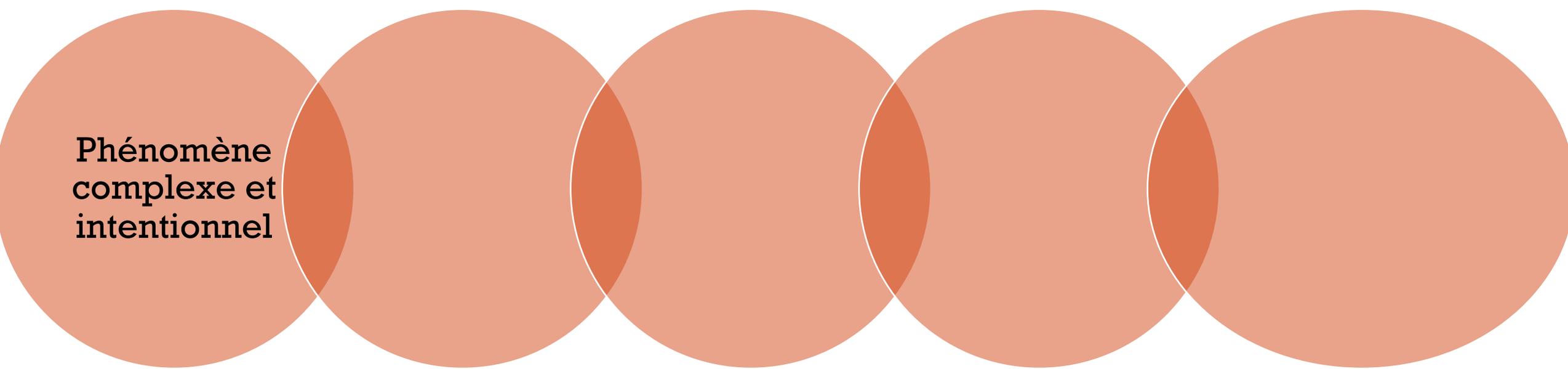
- Fraude à la carte
- Fraude à l'assurance
- Fraude sociale et fiscale
- **Blanchiment d'argent**



Entre 2% et 5% du PIB mondial *



FRAUDE = DÉFIS



Phénomène
complexe et
intentionnel



FRAUDE = DÉFIS

Phénomène
complexe et
intentionnel

Modèles de
détection de
fraude
difficiles à
évaluer



FRAUDE = DÉFIS

Phénomène
complexe et
intentionnel

Modèles de
détection de
fraude
difficiles à
évaluer

Evènement
rare



FRAUDE = DÉFIS

Phénomène
complexe et
intentionnel

Modèles de
détection de
fraude
difficiles à
évaluer

Evènement
rare

La fraude
évolue dans
le temps



FRAUDE = DÉFIS

Phénomène
complexe et
intentionnel

Modèles de
détection de
fraude
difficiles à
évaluer

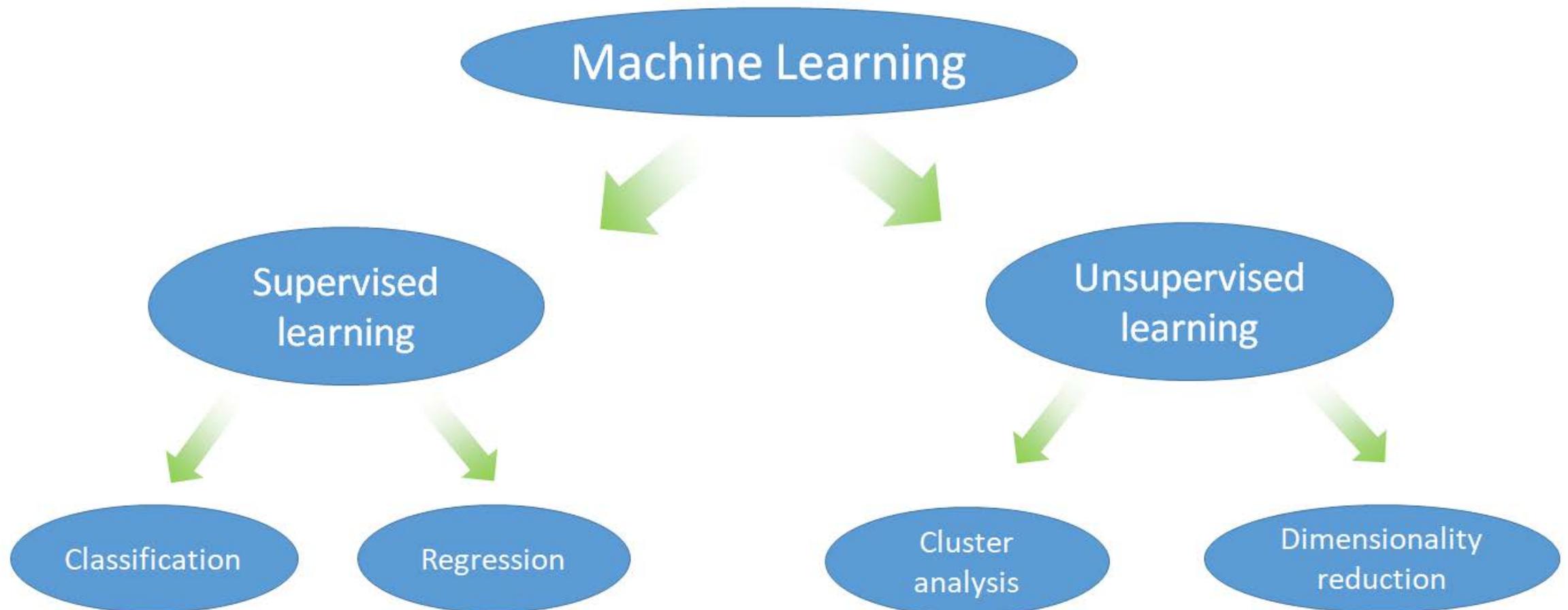
Evènement
rare

La fraude
évolue dans
le temps

Comprendre la
réglementation
en matière de
fraude et LCB



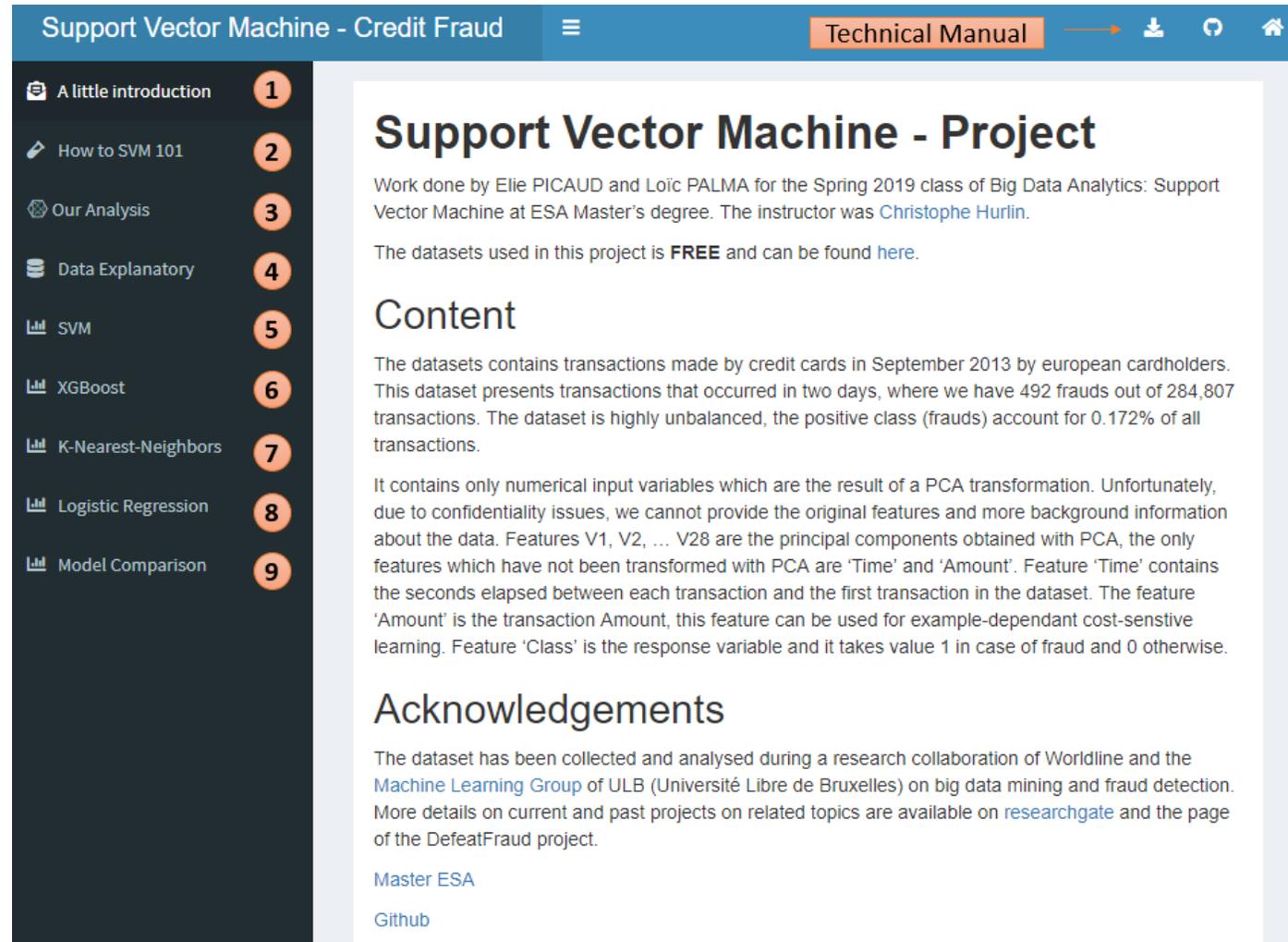
MÉTHODES DE DÉTECTION DE LA FRAUDE



MÉTHODES DE DÉTECTION DE LA FRAUDE

Projet M2 ESA

Détection de la fraude en utilisant différents modèles

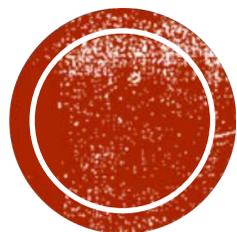


The screenshot shows a web browser displaying a technical manual for a project titled "Support Vector Machine - Credit Fraud". The browser's address bar shows the title and a "Technical Manual" tab. A dark sidebar on the left contains a table of contents with nine items, each with a numbered orange circle: "A little introduction" (1), "How to SVM 101" (2), "Our Analysis" (3), "Data Explanatory" (4), "SVM" (5), "XGBoost" (6), "K-Nearest-Neighbors" (7), "Logistic Regression" (8), and "Model Comparison" (9). The main content area has a white background and features a large heading "Support Vector Machine - Project". Below the heading is a paragraph of text: "Work done by Elie PICAUD and Loic PALMA for the Spring 2019 class of Big Data Analytics: Support Vector Machine at ESA Master's degree. The instructor was [Christophe Hurlin](#). The datasets used in this project is **FREE** and can be found [here](#)." This is followed by a "Content" section with a paragraph: "The datasets contains transactions made by credit cards in September 2013 by european cardholders. This dataset presents transactions that occurred in two days, where we have 492 frauds out of 284,807 transactions. The dataset is highly unbalanced, the positive class (frauds) account for 0.172% of all transactions." Below this is another paragraph: "It contains only numerical input variables which are the result of a PCA transformation. Unfortunately, due to confidentiality issues, we cannot provide the original features and more background information about the data. Features V1, V2, ... V28 are the principal components obtained with PCA, the only features which have not been transformed with PCA are 'Time' and 'Amount'. Feature 'Time' contains the seconds elapsed between each transaction and the first transaction in the dataset. The feature 'Amount' is the transaction Amount, this feature can be used for example-dependant cost-sensitive learning. Feature 'Class' is the response variable and it takes value 1 in case of fraud and 0 otherwise." The "Acknowledgements" section follows, stating: "The dataset has been collected and analysed during a research collaboration of Worldline and the [Machine Learning Group](#) of ULB (Université Libre de Bruxelles) on big data mining and fraud detection. More details on current and past projects on related topics are available on [researchgate](#) and the page of the DefeatFraud project." At the bottom of the page, there are two links: "Master ESA" and "Github".





APPEL À PROJETS DE RECHERCHE D'INITIATIVE ACADÉMIQUE



REDFLAG

DÉTECTION DE LA FRAUDE ET LUTTE CONTRE LE BLANCHIMENT D'ARGENT

L'OBJECTIF

- ✓ Identifier les **spécificités méthodologiques** inhérentes au domaine de la détection de la fraude et du blanchiment d'argent
- ✓ Proposer des **solutions techniques** innovantes en tirant parti d'une approche multidisciplinaire tout à fait originale.

Besoin de compétences en

data science et économétrie

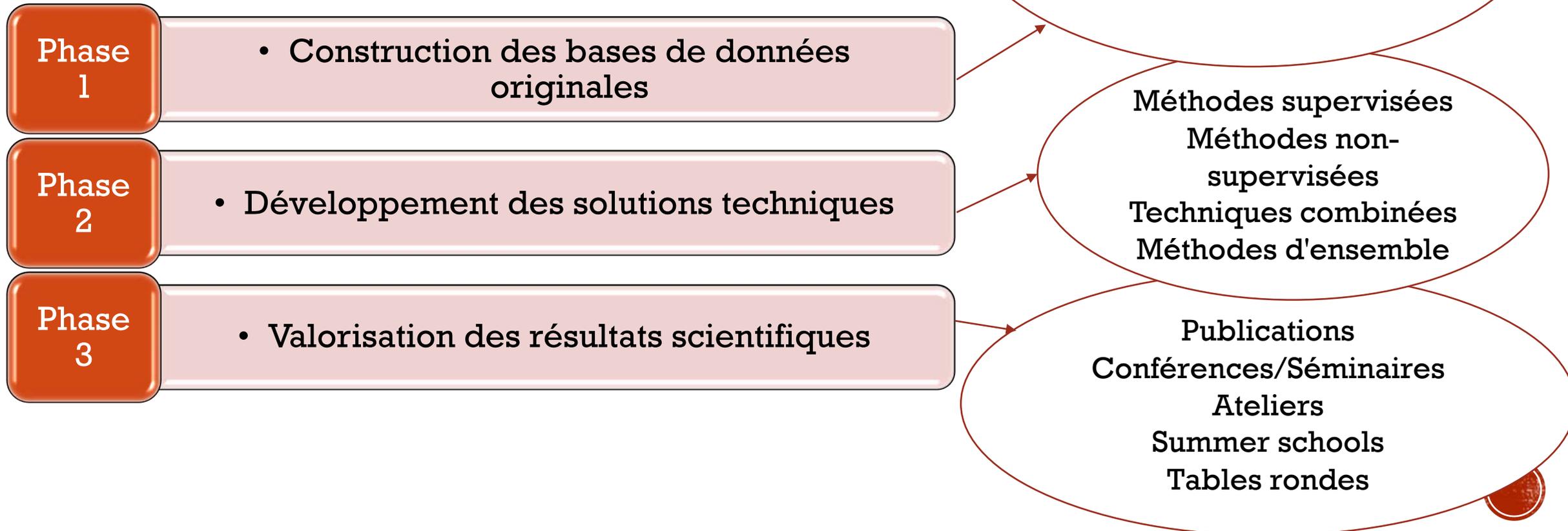
connaissance des modèles statistiques et de leurs propriétés

économie et droit

définitions, typologies, dispositifs de prévention, sanctions du blanchiment d'argent permettant de comprendre les motivations et les comportements stratégiques des fraudeurs



LE PROGRAMME DE TRAVAIL



L'INTERDISCIPLINARITÉ ET L'OUVERTURE À D'AUTRES COLLABORATIONS



Transdisciplinaire

- Economie
- Data science
- Droit
- Econométrie



D'intérêt

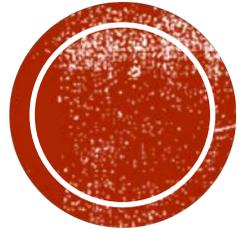
- Régional
- National
- International



Interlaboratoires

- LEO
- CRJP





DIPLÔMES D'UNIVERSITÉ

DIPLÔMES D'UNIVERSITÉ

- Offre de formation professionnelle continue (ou *executive education*) pour la rentrée 2019-2020 (5 certificats)

- Un certificat est une formation diplômante courte (9 jours de formation), centrée sur une problématique métier, et qui donne lieu à la délivrance d'un Diplôme d'Université (DU), à l'issue d'une évaluation.

EXECUTIVE
EDUCATION
L'ÉCOLE UNIVERSITAIRE
DE LA TRANSFORMATION DIGITALE
DES MÉTIERS DU DROIT ET DE L'ÉCONOMIE-GESTION



DIPLÔMES D'UNIVERSITÉ

Data Science pour la Détection de la Fraude Financière



Compliance et Lutte contre le Blanchiment d'Argent



DIPLÔMES D'UNIVERSITÉ

Data Science pour la Détection de la Fraude Financière

- **L'objectif général de la formation :**

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.

- **Le public ciblé :**

Data Scientists, ingénieurs statisticiens, chargés d'études statistiques, consultants, compliance officers, etc.

Compliance et Lutte contre le Blanchiment d'Argent

- **L'objectif général de la formation :**

former les participants aux **dispositifs légaux, réglementaires et techniques** de lutte contre le blanchiment d'argent.

- **Le public ciblé :**

Compliance officers, consultants, cadres des services de conformité et de gestion, etc.



DIPLÔMES D'UNIVERSITÉ

Data Science pour la Détection de la Fraude Financière

Module 1 :

- Détection de la fraude - éclairages juridiques (2 jours)

Module 2 :

- Gestion et traitement des données (1 jour)
- Méthodes non-supervisées pour la détection de la fraude (1 jour)

Module 3 :

- Méthodes supervisées pour la détection de la fraude (2 jours)

Module 4 :

- Evaluation des modèles de détection de la fraude (1 jour)
- Les réseaux sociaux et la fraude (1 jour)

Compliance et Lutte contre le Blanchiment d'Argent

Module 1 :

- Le cadre de la lutte contre le blanchiment d'argent (2 jours)

Module 2 :

- Les outils et les méthodes de la lutte contre le blanchiment d'argent (2 jours)

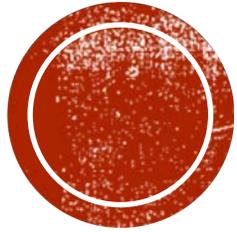
Module 3 :

- Les aspects procéduraux de la lutte contre le blanchiment d'argent (2 jours)

Module 4 :

- Les techniques d'investigation (2 jours)





MERCI

DENISA.BANULESCU-RADU @ UNIV-ORLEANS.FR