

Techniques d'apprentissage  
**IFT603-712**

**Présentation**

Par  
Pierre-Marc Jodoin

1

## Présentation

- **Professeur** : Pierre-Marc Jodoin
- **Courriel** : pierre-marc.jodoin@usherbrooke.ca
- **Page web perso** : [jodoin.github.io](https://jodoin.github.io)
- **Page web cours** : [jodoin.github.io/cours/ift603](https://jodoin.github.io/cours/ift603)
- **Local** : D4-1016-1, pavillon des sciences
- **Périodes de disponibilités** : [Lundi-vendredi de 9h00 à 17h00](#)

2

2

## Horaire

**Sujet à modification!**

- Cours magistraux
  - voir ici: [horaire.dinf.usherbrooke.ca](http://horaire.dinf.usherbrooke.ca)
- Travaux dirigés
  - Parfois en lab, voir vos courriels

3

3

## Évaluation IFT-603

- Examens
  - Intra 20%
  - Final 40%

} Seules les notes manuscrites non photocopiées seront admises
- Devoirs
  - 4 travaux pratiques de 10% chacun

4

4

## Évaluation IFT-712

- Examens
  - Intra 15%
  - Final 30%

} Seules les notes manuscrites non photocopiées seront admises
- Devoirs
  - 4 travaux pratiques de 7.5% chacun
- Projet de session
  - 25%

5

5

## Évaluation (suite)

IFT 712 => **projet de session**, voir plan de cours pour plus de détails

**[jodoin.github.io/cours/ift603/](http://jodoin.github.io/cours/ift603/)**

6

6

## Évaluation (suite)

- Les travaux pratiques se font en **python**.
- Correcteur et soutien technique
  - **Voir site web du cours**

7

7

## Évaluation (suite)

- Les examens se font **seul**.
- L'examen final portera sur toute la matière vue en classe
  - Date des examens :
    - **intra** à déterminer,
    - **final** à déterminer.
- Vous avez droit à des feuilles **manuscrites** pour toute documentation.
- Votre présence aux séances magistrales est fortement recommandée.
- Le cours est en **présentiel**.
- Une moyenne pondérée de moins de **35% aux examens entraîne automatiquement un échec.**

8

8

## Évaluation (suite)

- Les travaux pratiques se font **PAR ÉQUIPE DE TROIS ou QUATRE**.
  - Sinon **PÉNALITÉS**
  - **Pas d'équipe solo!**
- **Équité**
  - Les équipes seront formées **au hasard**
  - Par contre, les équipes peuvent être **scindées en cas de conflit**.
  - La correction (TP + examen) est aveugle
    - Donnez votre **login et votre matricule**

9

9

## Évaluation (suite)

- **IMPORTANT** : en plus de votre code, veuillez soumettre un fichier "**gitlab.txt**" dans lequel vous donnez le lien vers votre dépôt gitlab. Il est obligatoire d'utiliser gitLab (gitHub, Bitbucket, etc. sont interdits.)
- **IMPORTANT** : afin d'éviter toute discrimination, les équipes des travaux pratiques seront formées par l'enseignant.
- **IMPORTANT** : en plus de vos documents de travail, vous devez remplir, signer et joindre à votre travail le **formulaire d'intégrité** disponible sur le site web du cours.
- **IMPORTANT** : en plus de vos documents de travail, vous devez remplir et joindre à votre travail l'**agenda d'équipe** disponible sur le site web du cours.
- **OPTIONNEL** : un lien vers un sondage d'appréciation pour le travail d'équipe est disponible sur le site web du cours pour chaque travail.

10

10

## Évaluation (suite)

Une personne ayant peu contribué au travail d'équipe (**gitLab** et **agenda d'équipe**) s'expose à une **perte de points** et, à la limite, à avoir une note de **ZÉRO**.

**Rencontres hebdomadaires d'équipe obligatoires** (agenda d'équipe).

11

11

## Évaluation (suite)

- La remise du code et des exercices théoriques (lorsqu'il y en a) se fait par le système **TurninWeb**

(<http://turnin.dinf.usherbrooke.ca/>)

- Si vous avez des réponses manuscrites, vous devez les scanner.
- 10 points de pénalité par jour de retard
- 0 après 5 jours de retard
- Une erreur de remise **peut entraîner une note de zéro**.

- **PAS D'EXCEPTION!**

12

12

## Évaluation (suite)

- Avec le travail à distance, il est **obligatoire** d'utiliser un gestionnaire de code source « git ». Afin de simplifier les choses, veuillez utiliser le gitlab de l'UdeS:

[depot.dinf.usherbrooke.ca](https://depot.dinf.usherbrooke.ca)

- Pas de code envoyé par courriel!
- Une mauvaise utilisation de git pourra entraîner une **perte de points** pouvant aller jusqu'à **zéro**.
- **Aucune trace de code dans git = zéro**
- **Aucun code pertinent dans git = zéro**
- Vous ne connaissez pas git?

[www.tutorialspoint.com/git/index.htm](https://www.tutorialspoint.com/git/index.htm)

13

13

## Recommandations

- **N'attendez pas à la dernière minute pour faire les TP et le projet**
- Faites 100% des TP et non 50%-50%.
- Travaillez en équipe et non... côte à côte.
- Jamais une bonne idée de **plagier**
- Feedbacks en temps réel.
- **Rencontres hebdomadaires prévues au calendrier**
- Pénalité de 10% par jour de retard, à vous de ne **pas faire d'erreur** avec le système de remise « **turnin** »

14

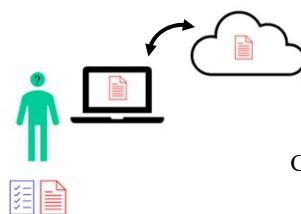
14

# NE PLAGIEZ PAS!

(voir dernière page du plan de cours)


15


## Qu'est-ce que de le plagiat?



Remise

Copier le travail d'Internet 

Copier le travail de ChatGPT 

Copier le travail d'Internet  
en **citant la source** 

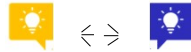
Copier le travail d'Internet  
avec approbation du prof

16




## Qu'est-ce que de la triche?

Travail individuel



S'aider mutuellement

Remettre le même travail 



Remise




17

## Qu'est-ce que de la triche?

IFTabc

IFTdef

Auto-plagiat 

Solution: consulter l'enseignant  
d'IFTdef



Remise



Remise

18

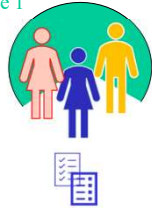
## Qu'est-ce que de la triche?

< >

Travail en groupe



Équipe 1



Équipe 2



Travail entre personnes  
d'équipes différentes

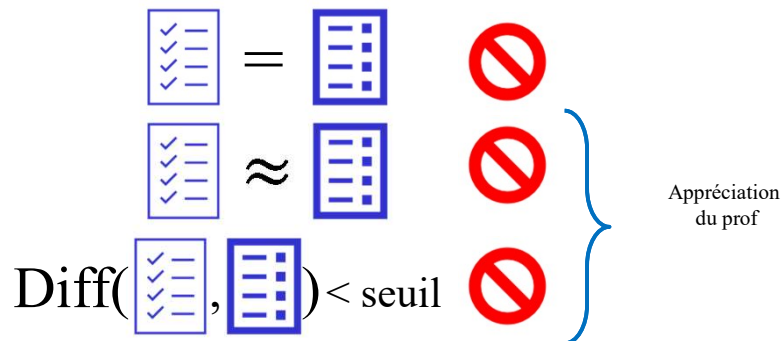


Remise



19

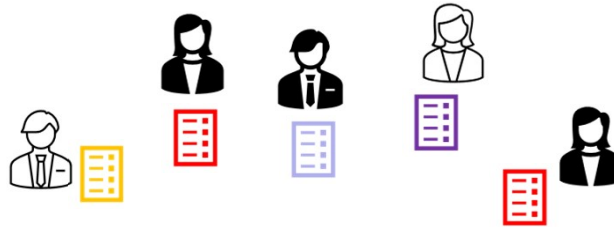
## C'est quoi Copier?



20

# Triche

Examen



!?

≈

Correcteur

21

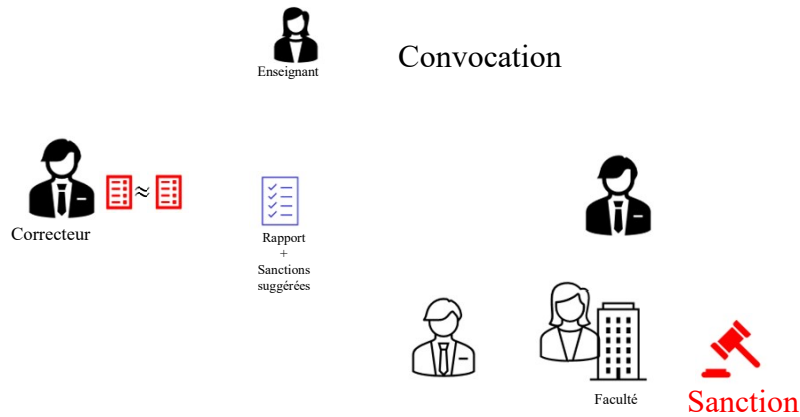
# Attention!!!

- Utilisation d'une partie de code ou texte qu'un de vos amis vous transmet  
○ **plagiat**
- Utilisation du code ou texte soutiré d'un travail disponible sur le web sans citer votre source  
○ **plagiat**
- Partage de réponses dans un examen, peu importe le moyen  
○ **tricherie**

si vous n'avez pas créé vous-même le code ou votre travail, vous êtes obligé de citer la référence, sinon il s'agit de plagiat. Votre travail doit être issu de votre **réflexion personnelle**.

22

# Procédure



23

# Sanctions

Peut varier en fonction des circonstances

Exemples types:

- 0 pour la partie plagiée
- ou
- 0 pour le travail ou l'examen

24

# Sanctions

De plus...

Une 1<sup>ère</sup> offense = inscription à une **liste noire facultaire**

Une 2<sup>e</sup> offense = conséquences plus graves (peut aller jusqu'à l'expulsion)

25

# Excuses non valides

- « J'ai travaillé fort »
- « Ça fonctionne ! »
- « C'est trop difficile »
- « Je ne savais pas »
- « C'est juste une petite partie »
- « Les versions ne sont pas identiques »
- « Je n'ai pas d'affinité avec mes collègues »

26

1<sup>er</sup> plagiat = *black list* facultaire

2<sup>e</sup> plagiat = expulsion

27

ChatGPT = plagiat

Code sur internet = plagiat

Code identique dans 2 travaux = plagiat

Travailler fort + plagiat = plagiat

Petit plagiat = plagiat

28

## À partir de maintenant

- Au cours de la 2<sup>e</sup> semaine de cours, **les équipes seront formées par l'enseignant.**
- Attention! Advenant un **conflit** dans une équipe (conflits de personnalité, méthodes de travail irréconciliables, manque d'ardeur au travail, etc.)
  - L'équipe pourra être scindée
  - Une personne pourrait être assignée à une autre équipe
- Si une personne est expulsée de 2 équipes, elle devra compléter la session **SEULE**.
- Une personne n'ayant **PAS contribué** à un travail s'expose à obtenir la **note de 0**.

29

29

## À partir de maintenant

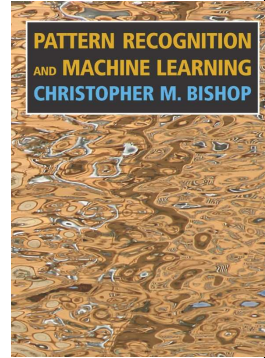
- Révision (ou apprentissage) de python (voir tutoriels sur le site du cours)
- Révision des bases en math (voir vidéos en ligne + chap.2 à 6 du livre de Deisenroth, Faisal et Ong “**Mathematics for machine learning**” )
- Visionnement du matériel de la première semaine (mise à niveau + concepts fondamentaux)

30

30

## Déroulement

- Site web du cours
  - [jodoin.github.io/cours/ift603/index.html](http://jodoin.github.io/cours/ift603/index.html)
- Livre obligatoire
  - *Pattern Recognition and Machine Learning*  
Christopher Bishop, Springer, 2007
- Où se procurer le livre?
  - Amazon  
[www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738](http://www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738)
  - Bibliothèque de science et de génie
  - Version gratuite en ligne!!  
[jodoin.github.io/cours/ift603/BishopBook.pdf](http://jodoin.github.io/cours/ift603/BishopBook.pdf)



31

31

## Introduction au cours

32

32



# Apprentissage automatique

- C'est une sous-discipline de l'intelligence artificielle en forte croissance.
- But de l'apprentissage automatique (*machine learning*)
  - Donner à un ordinateur la capacité d'apprendre à résoudre des problèmes par lui-même.
  - Découvrir les fondements théoriques de l'apprentissage
- Mathématiques-informatique appliqué
  - Informatique
  - Mathématique (stats, algèbre linéaire, probabilités, optimisation)

33

33

## Deux grands objectifs

- Apprendre les fondements **théoriques**
- Savoir les mettre en **pratique**

34

34

## Applications

- **Vision par ordinateur**
  - Reconnaissance de caractères
  - Localisation-reconnaissance de visages
  - Analyse d'images médicales
  - Reconnaissance de silhouettes humaines
- **Traitement automatique du langage**
  - système de questions-réponses (GPT-3, ChatGPT)
  - reconnaissance de la voix (Siri)
  - classification de documents (pourriels)
  - traduction automatique (google translate)
- **Robotique**
  - Conduite automatisée (détection d'obstacles, localisation de la route, détection de panneaux routiers, etc.)
- **Et bien bien bien d'autres**
  - Prédiction financière, recommandation d'achat (Amazon), etc.

35

35

## Liens avec d'autres cours

- **IFT 615 – Intelligence artificielle**
  - traite de plusieurs sous-disciplines de l'intelligence artificielle, pas seulement, l'apprentissage automatique
- **IFT 501 - Recherche d'information et forage de données**
  - Apprentissage non-supervisé (analyse de données, recommandation, données du web, etc.)
- **IFT 607 / 714 - Traitement automatique des langues naturelles**
- **IFT 608 / 702 - Planification en intelligence artificielle**
  - Planification et méthodes par renforcement.
- **ROP 317 / 630 – Programmation linéaire / Modèles de la recherche opérationnelle**
  - Optimisation
- **STT 418 – Programmation linéaire / Modèles de la recherche opérationnelle**
  - Statistiques et probabilités appliquées.
- **STT722 - Théorie de décision**
  - Analyse bayésienne et décision.

36

36

3 grands sujets au menu :

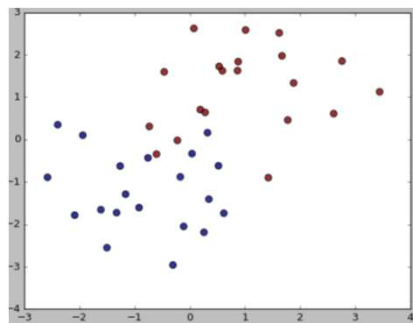
- Classification supervisée
- Régression supervisée
- Apprentissage non supervisé

37

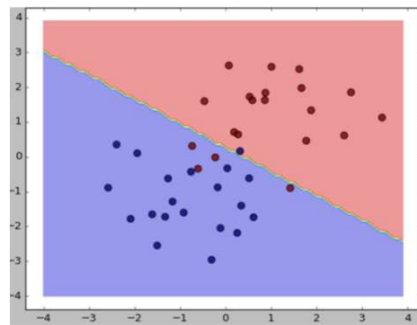
37

## Classification supervisée

Entraînement



À partir de données provenant de 2 classes  
● et ● (ici des “features” en 2D)



Le but est de trouver une fonction  $y()$  tel que

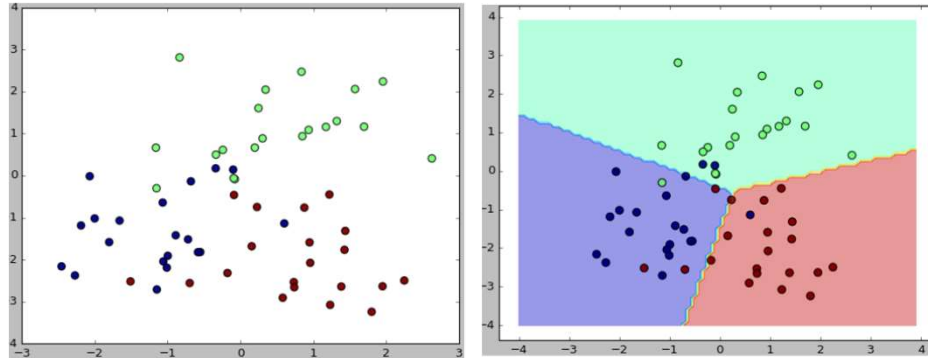
$y(\bullet)$  = classe 1

$y(\bullet)$  = classe 2

38

# Classification supervisée

Entraînement de plus de deux classes



Soit des données provenant de 3 classes ● et ● (ici des features en 2D)

Trouver une fonction  $y()$  tel que

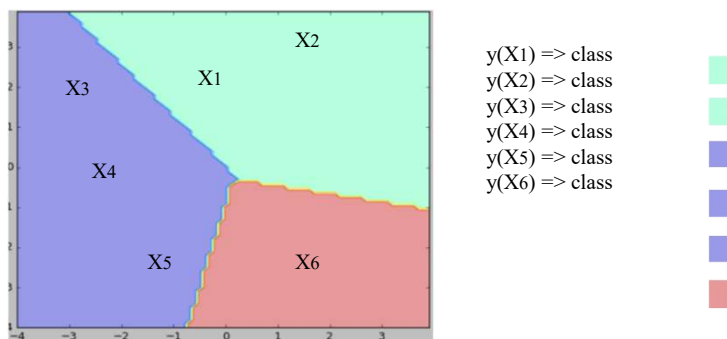
$y(\bullet) = \text{classe 1}$   
 $y(\bullet) = \text{classe 2}$   
 $y(\bullet) = \text{classe 3}$

39

# Classification supervisée

Une fois l'entraînement terminé, nous disposons d'un fonction  $y()$  pouvant convertir n'importe quel point 2D en une étiquette de classe.

La phase de test a pour but de prédire l'étiquette de classe de **nouvelles données  $X_i$  n'ayant jamais été observées**



$y(X1) \Rightarrow \text{class}$   
 $y(X2) \Rightarrow \text{class}$   
 $y(X3) \Rightarrow \text{class}$   
 $y(X4) \Rightarrow \text{class}$   
 $y(X5) \Rightarrow \text{class}$   
 $y(X6) \Rightarrow \text{class}$

40

## Classification supervisée (examples)

*Inria person dataset*

Positifs



Négatifs



41

## Classification supervisée (examples)

*Inria person dataset*

- 2 classes
- 20,252 images,
  - => 14,596 entraînement
  - => 5,656 test
- Chaque image est en RGB
  - => 64x128x3

Si on linéarise ces images, on peut les représenter à l'aide d'un vecteur de  $64 \times 128 \times 3 = 9,984$  dimensions.

42

## Classification supervisée (examples)

Partant de données étiquetées, (ici des images pour lesquelles on connaît l'étiquette de classe), l'apprentissage supervisé appliqué à la classification a pour objectif ultime de trouver une **fonction de classification** qui devrait permettre de bien classer de nouveaux exemples.

### *Inria Person dataset*

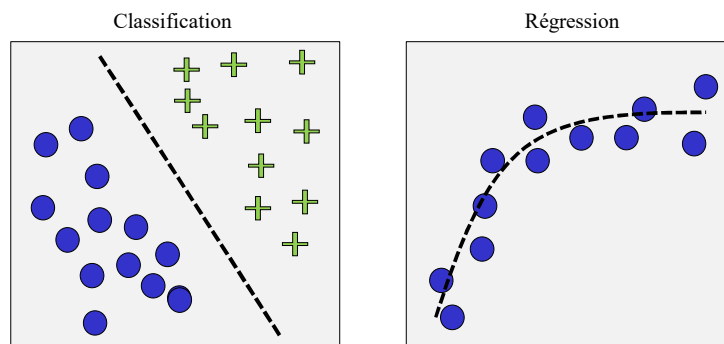
$y(\text{image d'une personne}) = \text{Person}$

$y(\text{image d'un paysage}) = \text{NotPerson}$

43

## Régression supervisée

Étant donné un ensemble d'entraînement, le but n'est pas de séparer les données de plusieurs classes mais de **retrouver le modèle ayant servi à générer les données**.

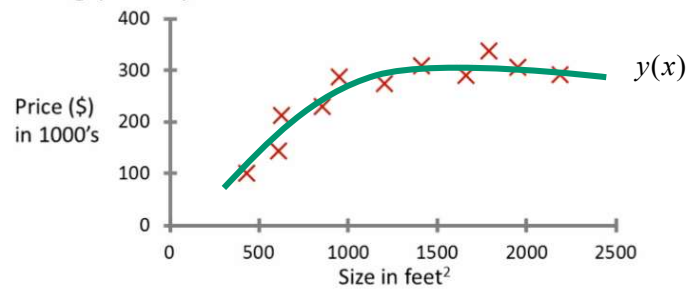


44

44

## Régression supervisée (exemples)

Housing price prediction.



Une fois  $y(x)$  connu, on peut estimer la valeur d'une maison étant donné une taille en  $\text{pi}^2$  jamais observée.

45

45

## Régression supervisée (exemples)

Prédictions boursières



Exemple, étant donné l'historique à la bourse d'une compagnie, quelle sera la valeur de son titre dans 1 semaine?

46

46

# Régression supervisée (exemples)

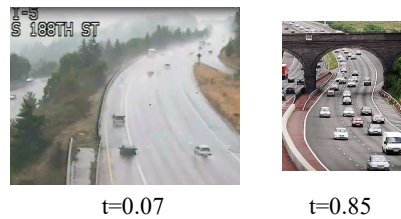
## Analyse de trafic routier

En entrée on a des images et en sortie la densité du trafic

Données d'entraînement



Données de test

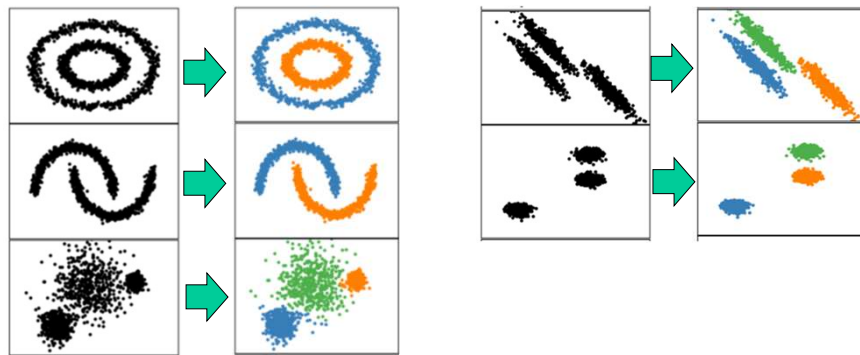


47

# Apprentissage

## Non-supervisé

Étant donné un ensemble de données **non-étiquetées**, grouper celles dont les caractéristiques sont similaires



On appelle également cette tâche “clustering”

48

48



# D'ici la semaine prochaine

## Révision

- Programmation Python 3.x (*Spyder*, *Pycharm*)
  - Apprenez à créer des environnements virtuels
  - Apprenez à installer des bibliothèques python avec « pip » ou « conda »
  - Configurez votre ordinateur!
  - Tutoriel python avec interface en web : [www.learnpython.org](http://www.learnpython.org)
  - Tutoriel python Stanford : [cs231n.github.io/python-numpy-tutorial](https://cs231n.github.io/python-numpy-tutorial)
  - Tutoriel python approfondi : [docs.python.org/3/tutorial](https://docs.python.org/3/tutorial)
- Se familiariser avec git ... **ça urge!**
- Se familiariser avec Linux
  - l'installer sur votre ordinateur
  - ou installer un VM Ubuntu (voir répertoire `Public/Logiciels/Ubuntu/`)
- Visionnement du matériel sur la mise à niveau
  - Dérivée
  - Dérivée partielle
  - Algèbre linéaire

49