

Techniques d'apprentissage  
**IFT603-712**

**Présentation**

Par  
Pierre-Marc Jodoin

**Présentation**

- **Professeur** : Pierre-Marc Jodoin
- **Courriel** : pierre-marc.jodoin@usherbrooke.ca
- **Page web** : <http://info.usherbrooke.ca/pmjodoin/>
- **Local** : D4-1016-1, pavillon des sciences
- **Période de disponibilités** : Jeudi-vendredi de 9h30 à 17h00

## Horaire

**Sujet à modification!**

- Cours magistraux
  - voir ici: [horaire.dinf.usherbrooke.ca](http://horaire.dinf.usherbrooke.ca)
- Travaux dirigés (*parfois, à voir en classe*)
  - Voir ici: <http://info.usherbrooke.ca/pmjodoin>

3

## Évaluation

- Les travaux pratiques se font en **python**.
- Correcteur et soutien technique
  - **Antoine Théberge**  
[antoine.theberge@usherbrooke.ca](mailto:antoine.theberge@usherbrooke.ca)

4

## 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**.

5

## Évaluation (suite)

- Les travaux pratiques se font **PAR ÉQUIPE DE DEUX ou TROIS**.
  - 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**

6

## 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 7 jours de retard
- Une erreur de remise **peut entraîner une note de zéro.**
- **PAS D'EXCEPTION!**

7

## 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](http://depot.dinf.usherbrooke.ca)

- Pas de code envoyé par courriel!
- Une mauvaise utilisation de git pourra entraîner une **perte de points** aux tp3, tp4 et pour le projet.
- Vous ne connaissez pas git?

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

8

## 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.
- Pénalité de 10% par jour de retard, à vous de ne **pas faire d'erreur** avec le système de remise « **turnin** »

9

# NE PLAGIEZ PAS!

(voir dernière page du plan de cours)

## Évaluation IFT-603

- Examens
  - Intra 20%
  - Final 40%

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

11

## Évaluation IFT-712

- Examens
  - Intra 15%
  - Final 30%

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

12

## Évaluation (suite)

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

[info.usherbrooke.ca/pmjodoin/cours/ift603/](http://info.usherbrooke.ca/pmjodoin/cours/ift603/)

13

## 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**.

14

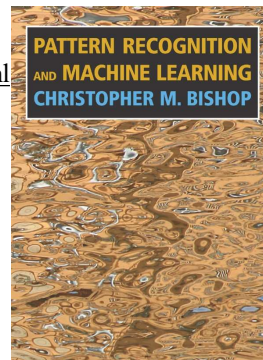
## À partir de maintenant

- 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)

15

## Déroulement

- Site web du cours
  - [info.usherbrooke.ca/pmjodoin/cours/ift603/index.html](http://info.usherbrooke.ca/pmjodoin/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!!  
[info.usherbrooke.ca/pmjodoin/cours/ift603/BishopBook.pdf](http://info.usherbrooke.ca/pmjodoin/cours/ift603/BishopBook.pdf)



16



# Introduction au cours

17

## 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 théorique de l'apprentissage
- Mathématiques-informatique appliqué
  - Informatique
  - Mathématique (stats, algèbre linéaire, probabilités, optimisation)

18

## Deux grands objectifs

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

19

## 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.

20

## 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.)
- **IFT702 - 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 decision.

21

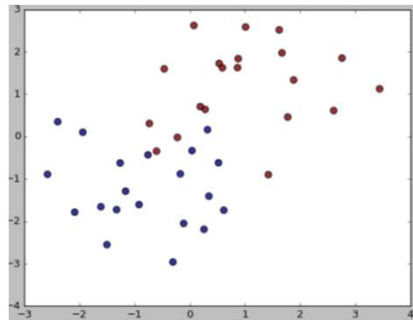
## 3 grands sujets au menu :

- Classification supervisée
- Régression supervisée
- Apprentissage non supervisé

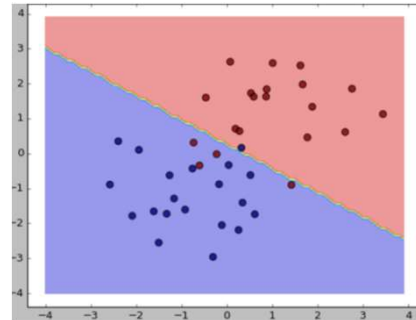
22

# Classification supervisée

Entraînement



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



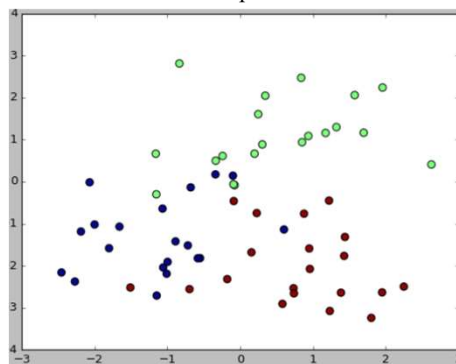
Le but est de trouver une fonction  $y()$  tel que

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

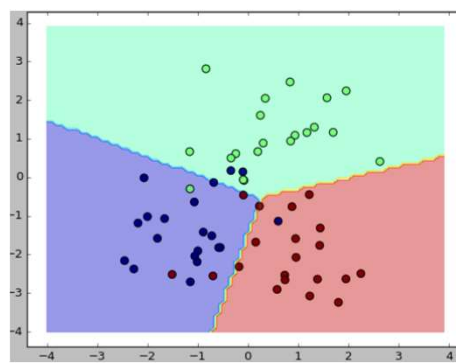
$y(\bullet) = \text{classe 2}$

# 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

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

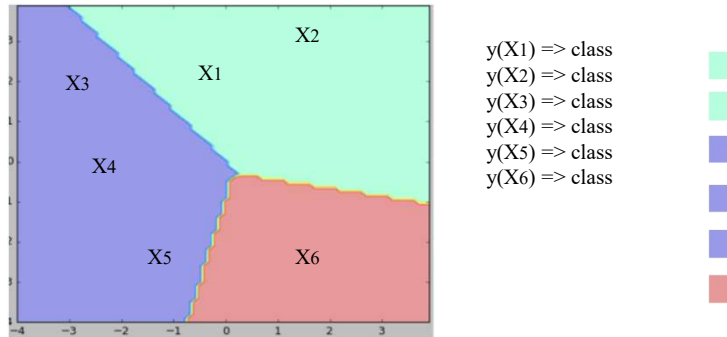
$y(\bullet) = \text{classe 2}$

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

# Classification supervisée

Une fois l'entraînement terminé, nous disposons d'une fonction  $y(\cdot)$  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**



# Classification supervisée (exemples)

*Inria person dataset*

Positifs



Négatifs



## Classification supervisée (examples)

*Inria person dataset*

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

Si on linéarise ces images, on peut les représenter à l'aide d'un vecteur de 64x128x3 = **9,984 dimensions**.

## Classification supervisée (examples)

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

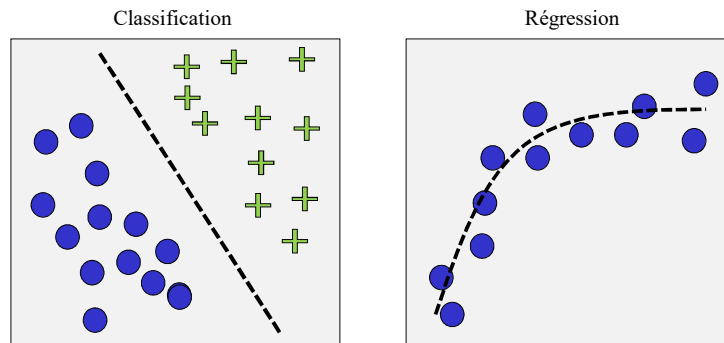
*Inria Person dataset*

$y(\text{img1}) = \text{Person}$

$y(\text{img2}) = \text{NotPerson}$

# Régression supervisée

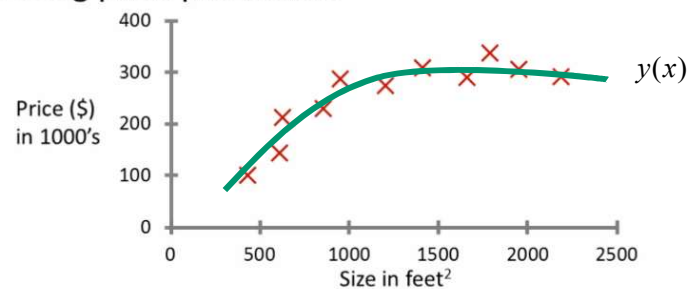
Étant donné un ensemble d'entraînement, le but n'est pas de séparer les données de plusieurs classes mais de **retrouver le modèle ayant servi à générer les données**.



29

## 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.

30

# 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?

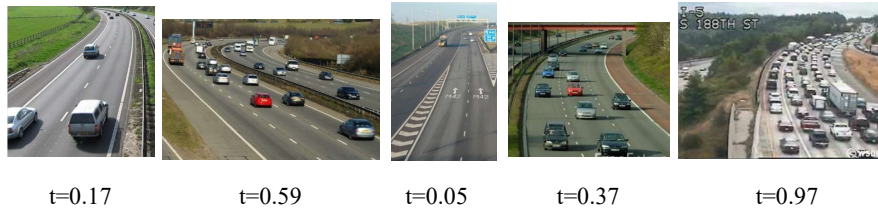
31

# Régression supervisée (exemples)

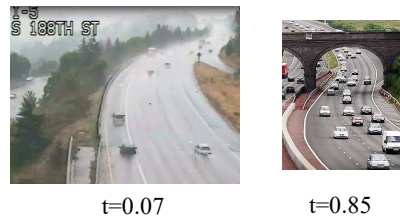
## Analyse de trafic routier

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

Données d'entraînement



Données de test

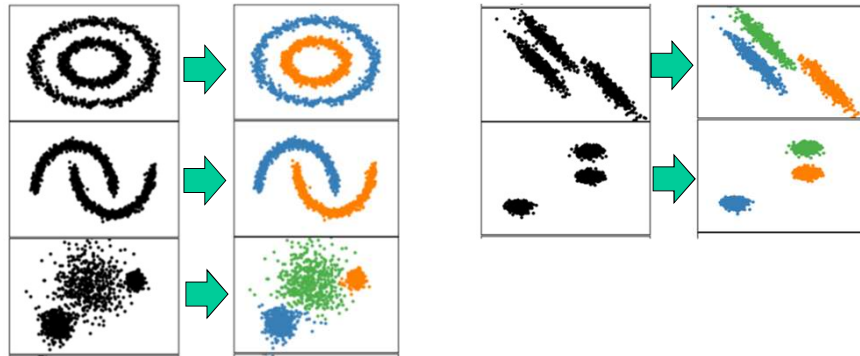




# Apprentissage

Non-supervisé

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



On appelle également cette tâche “**clustering**”

33

## 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 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

34