

Apprentissage automatique

Introduction à l'apprentissage automatique – GIF-4101 / GIF-7005

Professeur : Christian Gagné

Semaine 1



UNIVERSITÉ
LAVAL

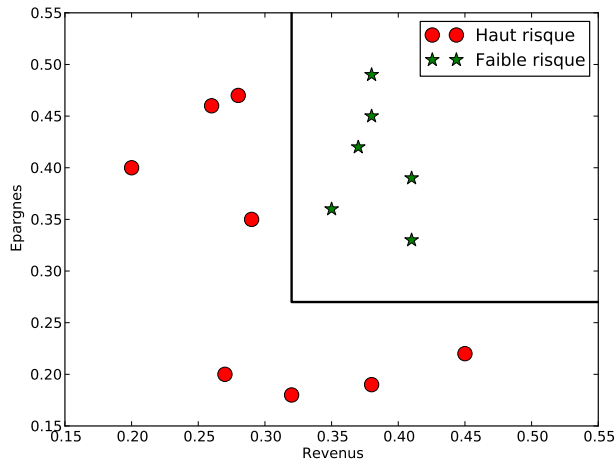
Pourquoi l'apprentissage automatique ?

- L'apprentissage automatique consiste à utiliser des ordinateurs pour **optimiser** un **modèle de traitement** de l'information selon certains **critères de performance** à partir d'**observations**, que ce soit des données-exemples ou des expériences passées.
- Lorsque l'on connaît le bon modèle de traitement à utiliser, pas besoin de faire de l'apprentissage !
- L'apprentissage automatique peut être utile lorsque :
 - On n'a pas d'expertise sur le problème (ex. robot navigant sur Mars)
 - On a une expertise, mais on ne sait pas comment l'expliquer (ex. reconnaissance de visages)
 - Les solutions au problème changent dans le temps (ex. routage de paquets)
 - Les solutions doivent être personnalisées (ex. biométrie)

Exemple

- Une entreprise en crédit doit estimer automatiquement le niveau de risque de ses clients
- Mesures disponibles : revenus du client (variable x_1) et épargnes du client (variable x_2)
- Banque de données sur des clients pré identifiés comme étant à haut risque (rouge) ou à faible risque (vert)

Exemple



Si $x_1 > 0,32$ et $x_2 > 0,27$ alors faible risque sinon haut risque

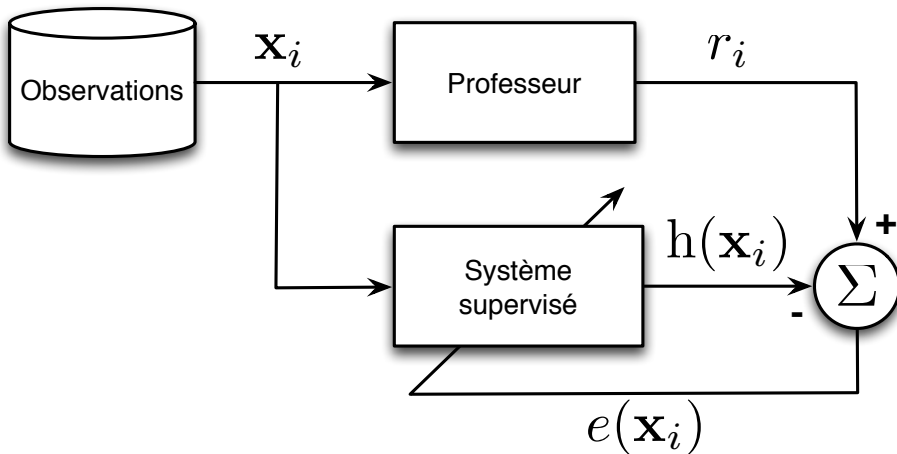
- Objectif : inférer un **modèle de traitement général** à partir d'**observations spécifiques**
 - Le modèle inféré doit être une approximation correcte et utile des observations
- Les observations sont peu coûteuses et disponibles en quantité suffisante ; la connaissance est coûteuse et rare
- Exemple : relier des transactions de consommateurs à leurs comportements de consommation
 - Suggestions d'items similaires sur Amazon (livre, musique), Netflix (cinéma), etc.

- Optimiser un modèle à partir d'observations selon un critère de performance
- **Vue statistique** : inférence à partir d'échantillons
- **Vue informatique** : construire des algorithmes et des représentations efficaces pour construire et évaluer les modèles
- **Vue ingénierie** : résoudre des problèmes sans devoir spécifier ou spécialiser manuellement les modèles

- Analyse d'un panier d'épicerie
 - $P(Y|X)$ est la probabilité qu'une personne achetant X achète également Y , où X et Y sont des produits ou services
- Exemple : $P(\text{croustilles}|\text{biere}) = 0,7$

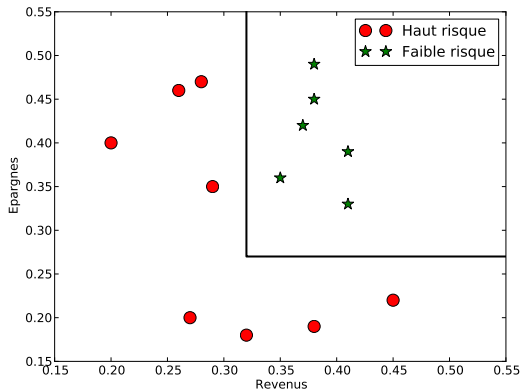
- Apprentissage supervisé
 - Objectif : apprendre une projection entre des observations X en entrée et des valeurs associées Y en sortie
- Modélisation mathématique
 - $y = h(x|\theta)$
 - $h(\cdot)$: fonction générale du modèle
 - θ : paramètres du modèle

Schématisation de l'apprentissage supervisé



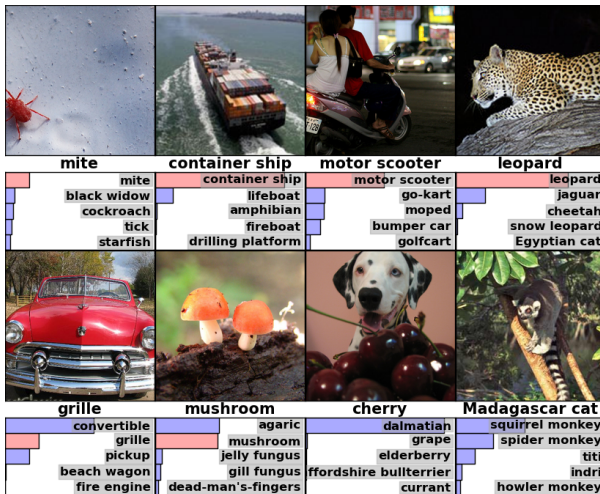
Classement

- Y est discret et correspond à des étiquettes de classes
- $h(\cdot)$ est une fonction discriminante



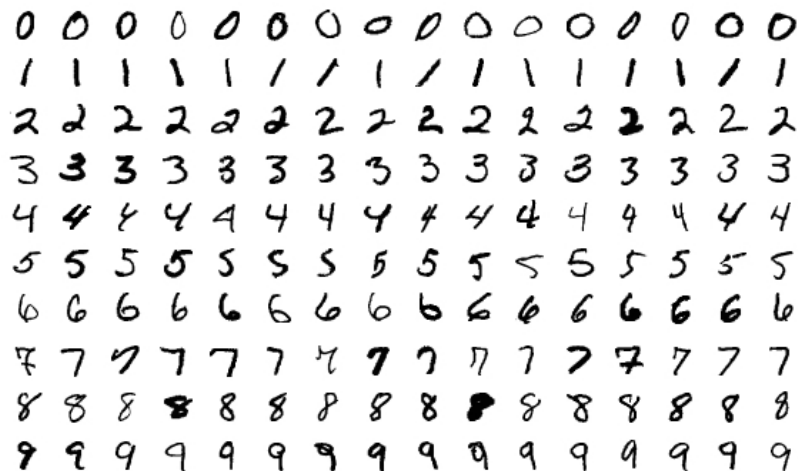
- Reconnaissance des formes
 - Reconnaissance d'objets : reconnaître les types d'objets présents dans une image malgré les variations des objets et des poses
 - Reconnaissance de caractères manuscrits : reconnaître malgré différents styles d'écriture
 - Reconnaissance de la parole : dépendance temporelle de l'information, utiliser des dictionnaires de mots/structures valides
- Traitement de la langue naturelle
- Aide au diagnostic médical
- Découverte de médicaments
- Biométrie
- Etc.

Reconnaissance d'objets



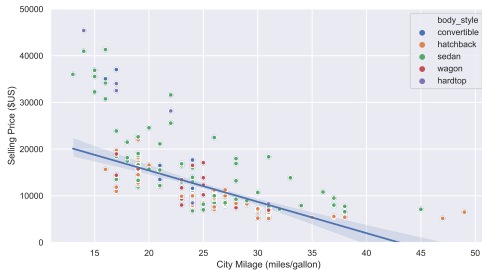
Tiré de A. Krizhevsky, I. Sutskever and G.E. Hinton. *Imagenet Classification with Deep Convolutional Neural Networks*. In *Advances in Neural Information Processing Systems*, 2012.

Reconnaissance de caractères



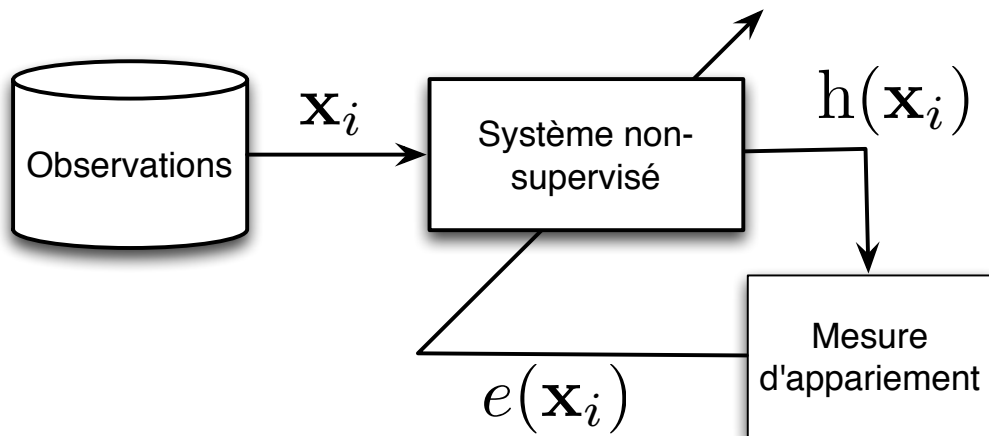
Régression

- Y est une valeur réelle
- $h(\cdot)$ est la fonction de régression
- Exemple : prédiction du prix de vente d'une voiture usagée à partir du kilométrage parcouru
 - Observations : kilométrage parcouru (x)
 - Prédiction : prix de vente (y)
- Applications en prédiction de valeurs
 - Finance et assurance
 - Phénomènes naturels (ex. météorologie)
 - Offre et demande
- Évaluation du risque et de l'incertitude



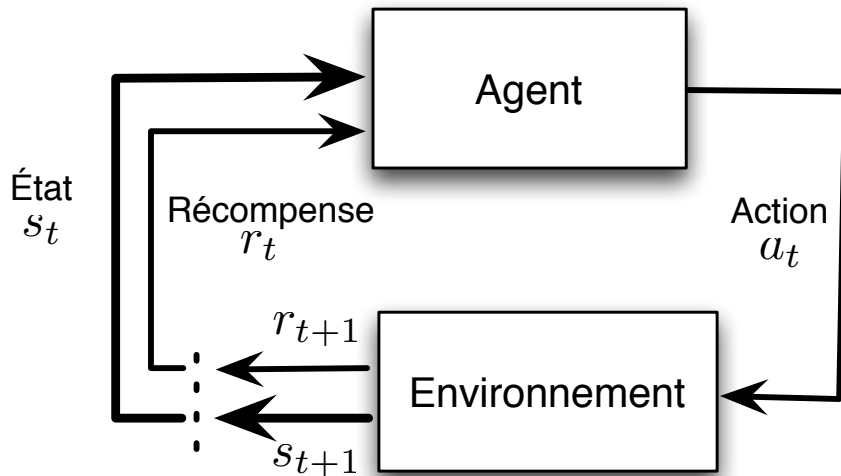
- Contrairement à l'apprentissage supervisé, il n'y a pas de valeurs de sortie
- Objectif : découvrir des régularités dans les observations
 - *Clustering* : découvrir des groupements d'observations similaires
- Applications
 - Segmenter les consommateurs dans les bases de données d'achats
 - Bio-informatique : découvrir des patrons dans l'ADN
 - Segmentation d'images : déterminer régions cohérentes d'images

Schématisation de l'apprentissage non supervisé



- Apprendre une politique : séquences d'actions
- L'apprentissage n'est pas supervisé, mais une récompense avec un délai est obtenue
- Problème d'assignation du crédit : quelles actions ont permis d'obtenir la récompense ?
- Applications
 - Jeux, avec un ou plusieurs participants
 - Robotique : navigation dans un environnement
 - Agents : prises de décisions

Schématisation de l'apprentissage par renforcement



- UCI Machine Learning Repository : <http://archive.ics.uci.edu/ml/>
- Kaggle
 - Compétitions : <https://www.kaggle.com/competitions>
 - Jeux de données : <https://www.kaggle.com/datasets>
- ImageNet : <http://www.image-net.org/>
- COCO (Common Objects in Context) : <http://cocodataset.org/>
- Données ouvertes
 - US : <https://www.data.gov/>
 - Europe : <http://data.europa.eu/euodp/fr/data/>
 - Canada : <http://ouvert.canada.ca/fr/donnees-ouvertes>
 - Québec : <https://www.donneesquebec.ca>