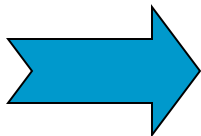




Apprentissage Automatique

Introduction

- Dans un système basé sur les connaissances la connaissance est le point central et essentiel ;
- Il est souvent difficile d'acquérir la connaissance d'un expert ;
- Existence d'experts non garantie (ex. prévision météo) ; par contre existence d'un historique !



Besoin d'automatiser le processus d'acquisition des connaissances

Pourquoi apprendre ?

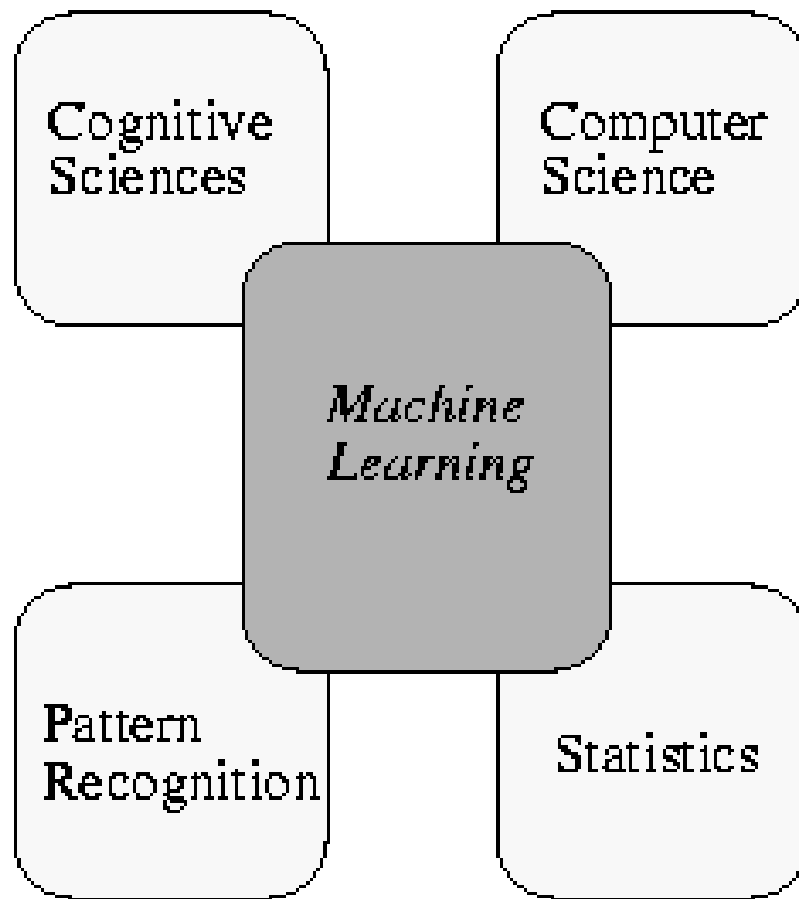
- L'apprentissage machine, c'est programmer les machines pour qu'elles optimisent un critère de performance, en utilisant des exemples ou des expériences passées.
- L'apprentissage est utile quand :
 - L'expertise humaine n'existe pas (se déplacer sur Mars),
 - Les humains sont incapables d'expliquer leur expertise (reconnaissance de la parole)
 - La solution change dynamiquement (routage dans un réseau)
 - La solution a besoin d'être adaptée à des cas particuliers (user biometrics)

«Apprentissage machine» : de quoi s'agit-il ?

- Apprendre des modèles généraux à partir de données, d'exemples, ...
- Les données sont peu coûteuses et abondantes (data warehouses, data marts) ; la connaissance est coûteuse et rare.
- Exemple :

Les gens qui ont acheté “Da Vinci Code” ont aussi acheté “The Five People You Meet in Heaven” (www.amazon.com)
- Construire des modèles qui soient de *bonnes et utiles approximations* des données.

Apprentissage



- Rôle des Statistiques : Inférer à partir d'un échantillon
- Rôle de l'informatique : algorithmes efficaces pour,
 - Solutionner un problème d'optimisation
 - Représenter et évaluer le modèle pour l'inférence

Définitions

- « a learning system uses sample **data** to generate an updated basis for **improved performance** on subsequent data, and expresses the new basis in **intelligible** symbolic form » [Donald Michie, 1991] ;
- « learning denotes changes in the system that are adaptative in the sense that they enable the system to do the same task or tasks drawn from the same population more effectively the next time » [Herbert Simon, 1983]

Applications

- Vente : Analyse de panier, ...
- Finance : Études de crédit, détection de fraudes
- Manufacture : Optimisation, troubleshooting
- Médecine : Diagnostic
- Télécommunications : Optimisation de la qualité de service
- Bioinformatique : Motifs, alignement
- Fouille dans le Web : Moteurs de recherche
- Sécurité : Cyber forensics
- ...

Choisir une approche d'apprentissage : contraintes

- Objectifs
 - Améliorer les performances ;
 - Acquérir de nouvelles connaissances ;
- Disponibilité d'une connaissance du domaine
- Disponibilité de données
 - incrémental vs. batch
- Caractéristiques des données
 - Prédicats, propositions, etc.

Choisir une approche d'apprentissage : contraintes

- Mécanisme de recherche
 - Model-driven (Top-down)
 - Data-driven (Bottom-up)
 - Autres
- Raisonnement
 - Induction
 - Abduction
 - Dédution

Tâches d'apprentissage

- Apprentissage supervisé
 - Knowledge poor
 - Knowledge intensive
- Résolution de problèmes
 - Apprendre des conditions d'application d'opérateurs
 - Apprendre à résoudre d'autres problèmes
- Apprentissage non supervisé
 - Pas d'oracle
- Découverte scientifique
 - Découverte de lois mathématiques, etc.

Comment évaluer un apprentissage

- Mathématiques
- Empirique
 - Populaire
 - Nécessite beaucoup de données
- Psychologie
 - Comparer à un modèle humain

Différentes techniques d'apprentissage

Classification des techniques en fonction du but principal.



L'apprentissage synthétique vise principalement à créer des connaissances nouvelles ou meilleures --> induction



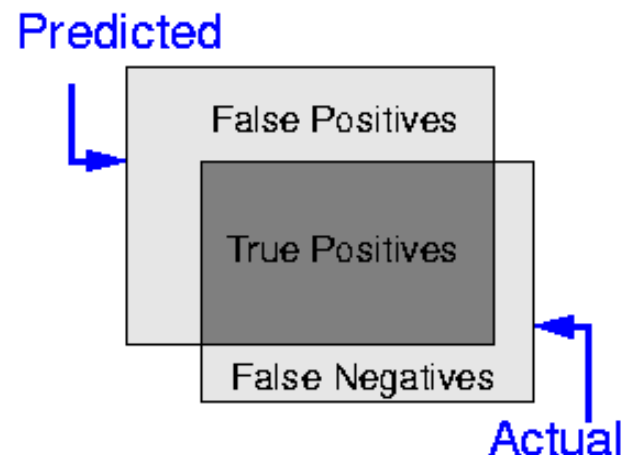
L'apprentissage analytique vise plutôt à mettre les connaissances sous une forme plus appropriée --> déduction

induction : processus de formation d'hypothèses - les prémisses - qui impliquent des conséquents donnés,

déduction : processus de dérivation de conséquents d'après des prémisses données.

Apprentissage supervisé sans connaissances du domaine (knowledge poor)

- À partir d'un ensemble d'exemples, on induit une hypothèse (description du concept) ;
- Mesures de performances typiques :
 - Capacité prédictive
 - Intelligibilité
 - Vitesse d'apprentissage et de test
 - Besoins en stockage



Apprentissage supervisé sans connaissances du domaine (knowledge poor) : exemples

- Arbres de décision (TDIDT)
- Induction de règles
- Espace des versions
- Algorithmes génétiques
- Réseaux de neurones
- Réseaux Bayésien (?)
- Etc.

Applications

- Association
- Apprentissage supervisé
 - Classification
 - Régression
- Apprentissage non supervisé
- Apprentissage par renforcement

Apprendre des associations

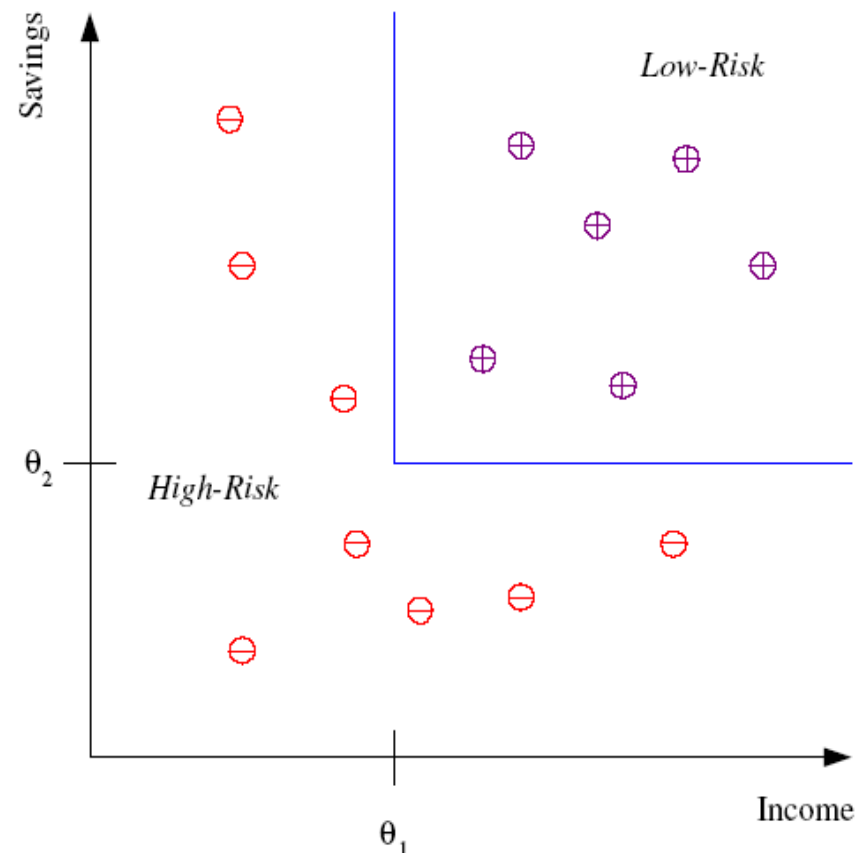
- Analyse de panier :

$P(Y | X)$ probabilité qu'un client qui achète X achète aussi Y où X et Y sont des produits/services.

Exemple: $P(\text{chips} | \text{bière}) = 0.7$

Classification

- Exemple: Analyse de crédit
- Différentier entre clients à **low-risk** et ceux à **high-risk** à partir de leurs revenus et épargnes.



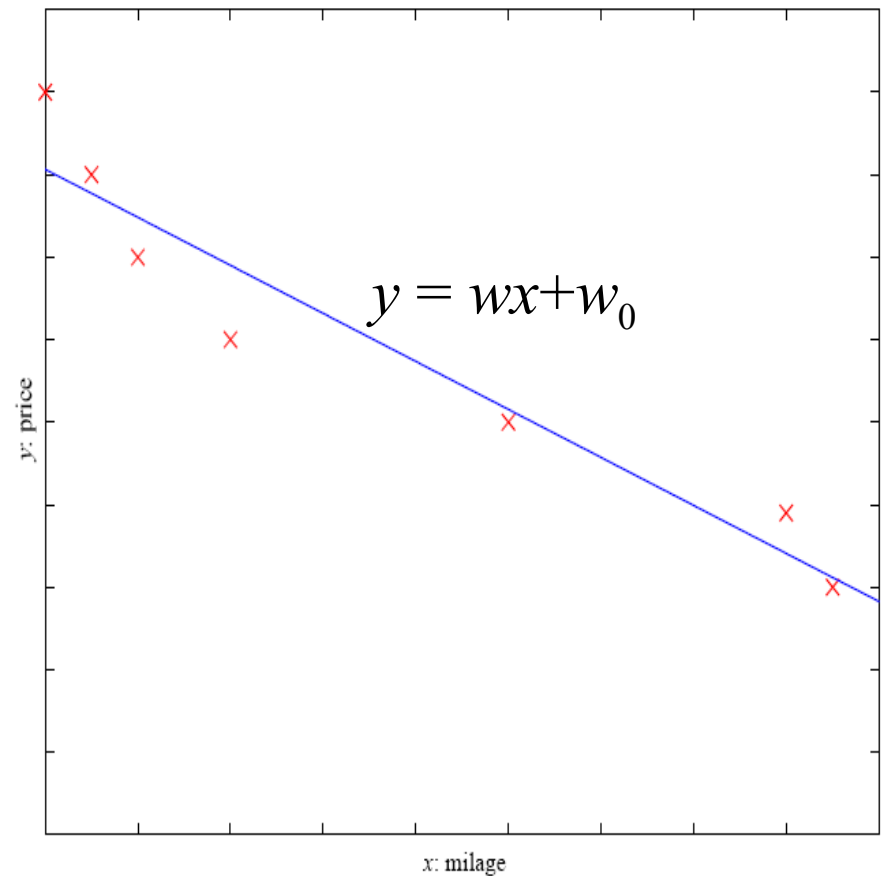
Discriminant : IF $income > \theta_1$ AND $savings > \theta_2$
 THEN **low-risk** ELSE **high-risk**

Classification : Applications

- Reconnaissance de formes :
 - Reconnaissance de visages
 - Reconnaissance de caractères
 - Reconnaissance de la parole
- Diagnostic médical : des symptômes aux maux
- ...

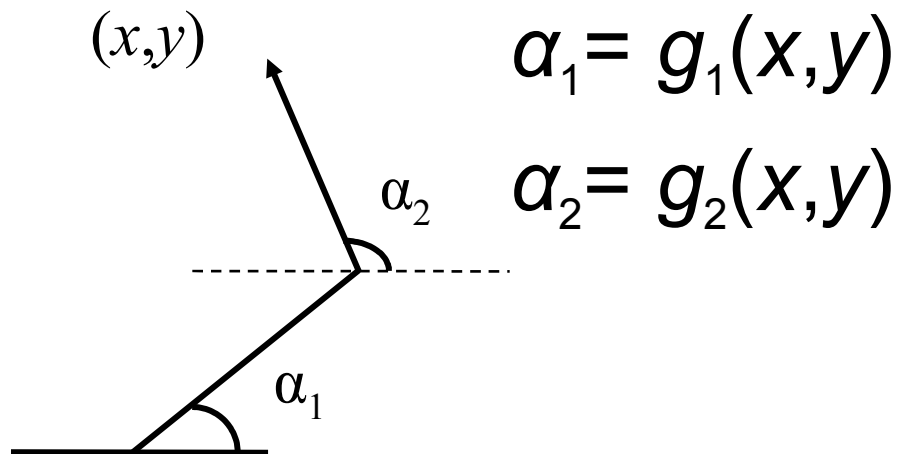
Régression

- Exemple : Prix d'une auto usagée
- x : attributs de l'auto
- y : prix
- $y = g(x | \theta)$
- $g(\cdot)$ modèle,
- θ paramètres



Régression : Applications

- Conduite d'un robot : Angle d'une roue motrice d'un robot (CMU NavLab)
- Kinématique d'un bras robotisé



- Machine à torréfier le café

Apprentissage supervisé : Usage

- Prédiction de cas futurs : Utiliser une règle pour prédire la « classe » (output) de cas futurs
- Extraction de connaissances : Une règle est simple à comprendre
- Compression : Simplicité d'une règle par rapport aux données qu'elle explique
- Détection d' «Outlier» : Exceptions qui ne sont pas couverts par la règle, e.g., une fraude.

Apprentissage non supervisé

- Apprendre des régularités dans les données
- Pas d'output
- Clustering : Regrouper des instances similaires
- Exemples d'applications :
 - Segmentation de clients
 - Compression d'images : Quantification de couleurs
 - Bioinformatique : Apprendre des motifs (séquences d'acides aminés dans des protéines, ...)

Apprentissage par renforcement

- Apprendre une politique : Une séquence d'outputs (ou d'actions)
- Le but est la réalisation d'une séquence gagnante : récompense retardée
- Jeux
- Robot dans un labyrinthe
- Agents multiples (robots soccer), observation partielle (incertitude), ...

Ressources : Données

- UCI Repository : <http://archive.ics.uci.edu/ml/>
- UCI KDD Archive :
<http://kdd.ics.uci.edu/summary.data.application.html>
- Statlib : <http://lib.stat.cmu.edu/>
- Delve : <http://www.cs.utoronto.ca/~delve/>

Ressources : Journaux

- Journal of Machine Learning Research
www.jmlr.org
- Machine Learning
- Neural Computation
- Neural Networks
- IEEE Transactions on Neural Networks
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- Annals of Statistics
- Journal of the American Statistical Association
- ...

Ressources : Conférences

- ✓ International Conference on Machine Learning (ICML), ICML11 : <http://www.ourglocal.com/url/?url=www.icml-2011.org>
- ✓ European Conference on Machine Learning (ECML), ECML11 : <http://www.ecmlpkdd2011.org/>
- ✓ Neural Information Processing Systems (NIPS), NIPS11 : <http://nips.cc/Conferences/2011/>
- ✓ Uncertainty in Artificial Intelligence (UAI), UAI11 : <http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=13866©ownerid=14179>
- ✓ International Joint Conference on Artificial Intelligence (IJCAI), IJCAI11 : <http://ijcai-11.iia.csic.es/>
- ✓ International Conference on Neural Networks (Europe), ICANN11 : <http://www.ourglocal.com/event/?eventid=6764>
- ✓ ...