

IA et science des données

Cours 11 – mardi 12 avril 2022

Christophe Marsala
Vincent Guigue

Sorbonne Université

LU3IN026 - 2021-2022

Programme du jour

Le projet

L'intelligence artificielle

IA, apprentissage et science des données

Pour conclure...

1 – Le projet –

Objectif du projet : analyse des données fournies

- Résultat attendu
 - définition d'un certain nombre de problématiques et leur résolution
 - au moins un problème d'apprentissage supervisé et un autre non-supervisé
- Rendu attendu :
 - un unique notebook éventuellement complété par un package avec vos fonctions utilisées
 - un poster (électronique) expliquant de façon synthétique en une page les différentes expériences réalisées et leurs résultats
- Calendrier (groupe de TDTME)
 - séance 11 libre : temps prévu pour travailler sur le projet (en partie)
 - **lundi 9 mai, à 19h** : dernière limite pour rendre le projet (Moodle)
 - Après la semaine d'examen (connu sous peu) : soutenances

Marsala & Guigue – 2022

LU3IN026 – cours 11 – 3

1 – Le projet –

Conseils ?

- Choisir une problématique qui vous intéresse : poser une question
 - existe-t-il des groupes d'aliments ?
 - peut-on prédire la qualité des données (bonne, moyenne, mauvaise) à partir des mesures ?
 - prédire le type d'aliment ?
 - ...
- Choisir un protocole d'expérimentation
 - Train vs Test, Validation croisée
 - indices d'évaluation
- Conclure des résultats obtenus

Marsala & Guigue – 2022

LU3IN026 – cours 11 – 5

1 – Le projet –

Les données : Agribalyse

- Evaluation environnementale des produits agricoles et alimentaires
 - <https://agribalyse.ademe.fr/>
 - pensez à lire la documentation !!
 - les tables peuvent être croisées
- Comment procéder ?
 - choisir une classe à prédire ?
 - on peut créer une classe binaire à étudier en coupant en 2 un univers numérique...
 - trouver des clusters ?
- Algorithmes de base
 - supervisé : kppv, perceptron (Rosenblatt, biais), arbres de décision
 - non-supervisé : clustering hiérarchique, K -moyennes

Marsala & Guigue – 2022

LU3IN026 – cours 11 – 4

2 – L'intelligence artificielle – une définition

Qu'est-ce que l'Intelligence artificielle ?



Turing



Shannon



Simon / Newell



McCarthy



Minsky

Une définition de l'IA

Science dont le but est de faire faire par la machine des activités que l'homme accomplit en utilisant son intelligence.

- Étude des activités intellectuelles pour lesquelles aucune méthode n'est, à priori, connue. (Laurière, 1986)
- \Rightarrow résoudre des problèmes pour lesquels il n'y a pas d'algorithme connu qui les résolve en temps raisonnable

Marsala & Guigue – 2022

LU3IN026 – cours 11 – 6

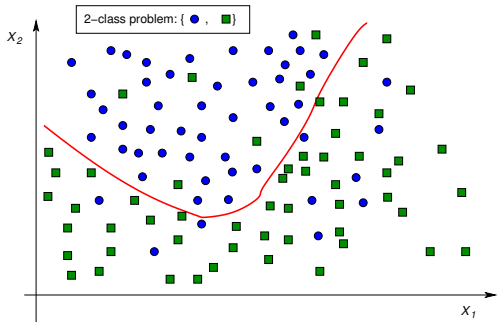
Qu'est-ce que l'Intelligence artificielle ?

- ▶ Approches de résolution : informatique heuristique
 - choix guidés par des **heuristiques**
 - représentation et simplification de l'espace des solutions
- ▶ Quelques exemples
 - reconnaissance de forme
 - résumé de texte, traduction, reconnaissance de style d'auteur
 - prise de décision (diagnostic médical, résolution de pannes,...)
 - modélisation de systèmes complexes (systèmes biomimétiques,...)
 - jeux : échecs, go, gestion de PNJ, moteur de jeu (IA)

IA, apprentissage et science des données

- ▶ Apprentissage artificiel
 - doter un ordinateur de **capacités d'apprentissage**
 - se servir des données du passé pour les données futures
 - différents cadres d'application :
 - mise au point (paramétrage de systèmes)
 - systèmes autonomes
 - Sciences de données
 - extraire des connaissances à partir de masses de données
 - l'apprentissage en est une étape
- ▶ Apprentissage : construire un classifieur $f : \mathbb{R}^d \rightarrow Y$
 - base d'apprentissage : ensemble de n exemples $\mathbf{X} \in \mathbb{R}^{n \times d}$
 - exemples (ou cas, ou individus) : $\mathbf{x}_1, \dots, \mathbf{x}_n$
 - chaque \mathbf{x}_i est décrit par d variables $x_{i,1}, \dots, x_{i,d}$
 - apprentissage supervisé : chaque \mathbf{x}_i est associé à un label $y_i \in Y$

Apprentissage supervisé : séparateur des classes

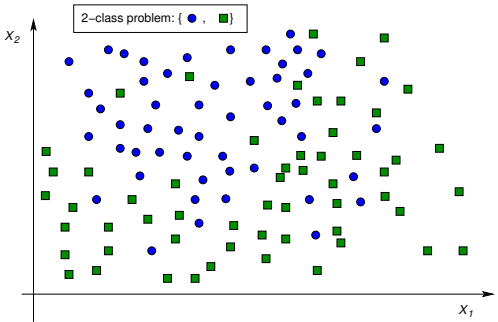


- ▶ Il existe une **frontière** qui sépare les classes
 - inconnue et doit être trouvée
 - ici : \mathbf{x} est un point (x_1, x_2) de $X_1 \times X_2$

Quelques dates clés

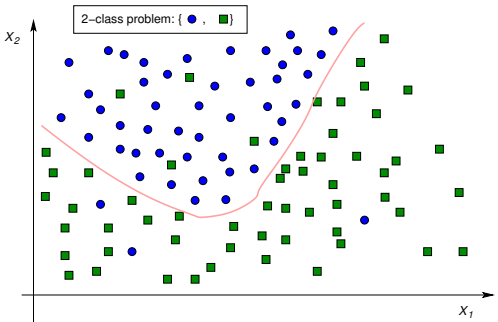
- ▶ 1949 : C. Shannon propose un algorithme pour le jeu d'échecs
- ▶ 1950 : Alan Turing : "Computing machinery and intelligence"
 - le jeu de l'imitation (test de Turing)
- ▶ 1956 : **Conférence de Dartmouth College**
 - Marvin Minsky, John McCarthy, Herbert Simon,...
 - "Logic theorist" : démonstrateur de théorèmes
 - but : simuler le raisonnement humain pour résoudre des problèmes
- ▶ ...
- ▶ 1996 : Deep Blue (IBM) bat Garri Kasparov (jeu d'échecs)
- ▶ 2005 : succès au DARPA grand challenge
 - Université de Stanford (Sebastian Thrun) : voiture sans pilote
- ▶ 2011 : victoire de l'IA Watson (IBM) au jeu Jeopardy
- ▶ 2016 : Alpha Go (Google Deepmind) bat Lee Sedol (4-1) au Go

Apprentissage supervisé



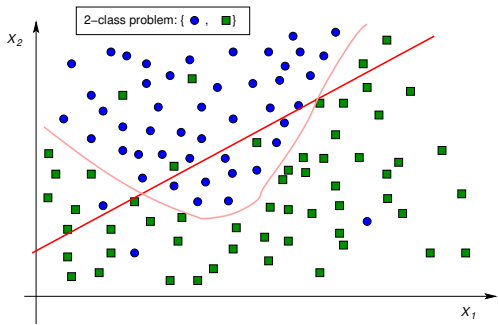
- ▶ Étant donné une base d'apprentissage
 - descriptions + classes
 - ici : \mathbf{x} est un point (x_1, x_2) de $X_1 \times X_2$

Apprentissage supervisé : comment séparer les classes



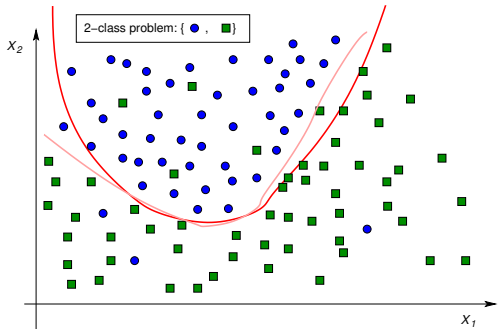
- ▶ Un classifieur est un estimateur de la frontière
 - **classifieur** : fonction f telle que $y = f(\mathbf{x})$
 - ici : \mathbf{x} est un point (x_1, x_2) de $X_1 \times X_2$ et y est la classe

Apprentissage supervisé : séparateur des classes



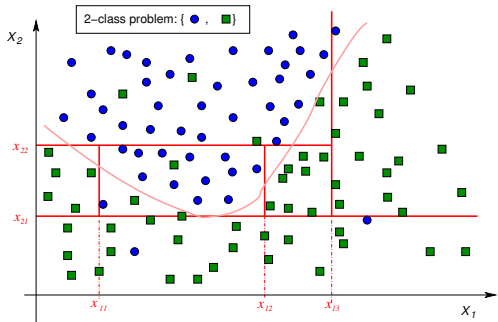
- Une possibilité : trouver a et b tels que
 - **classifieur linéaire** : fonction f telle que $y = ax + b$
 - ici : x est un point (x_1, x_2) de $X_1 \times X_2$ et y est la classe

Apprentissage supervisé : séparateur des classes



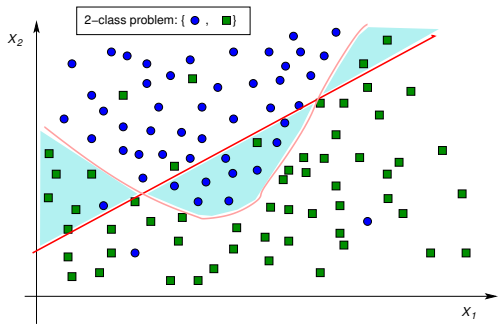
- Une autre possibilité : plus de paramètres
 - **classifieur quadratique** : fonction f telle que $y = ax^2 + bx + c$
 - ici : x est un point (x_1, x_2) de $X_1 \times X_2$ et y est la classe

Apprentissage supervisé : séparateur des classes



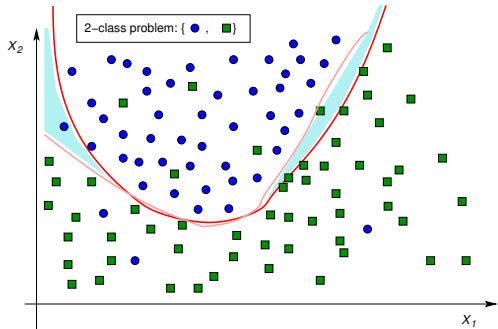
- Encore une autre possibilité : trouver des seuils de coupures x_{ij}
 - **arbre de décision** : fonction f composée de coupure perpendiculaires aux axes

Apprentissage supervisé : séparateur des classes



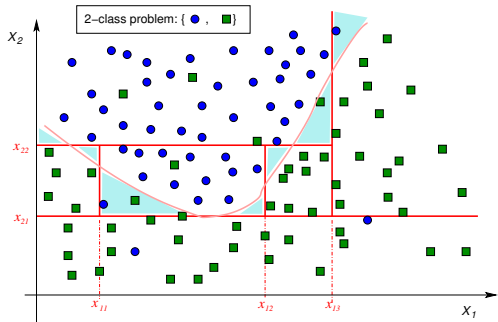
- Des régions peuvent contenir des exemples mal-classés
 - **classifieur linéaire** : fonction f telle que $y = ax + b$
 - ici : x est un point (x_1, x_2) de $X_1 \times X_2$ et y est la classe

Apprentissage supervisé : séparateur des classes



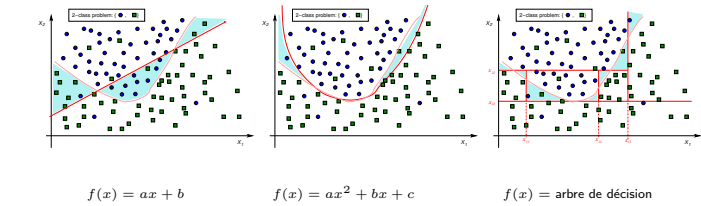
- Les régions avec les exemples mal-classés sont plus réduites
 - **classifieur quadratique** : fonction f telle que $y = ax^2 + bx + c$
 - ici : x est un point (x_1, x_2) de $X_1 \times X_2$ et y est la classe

Apprentissage supervisé : séparateur des classes



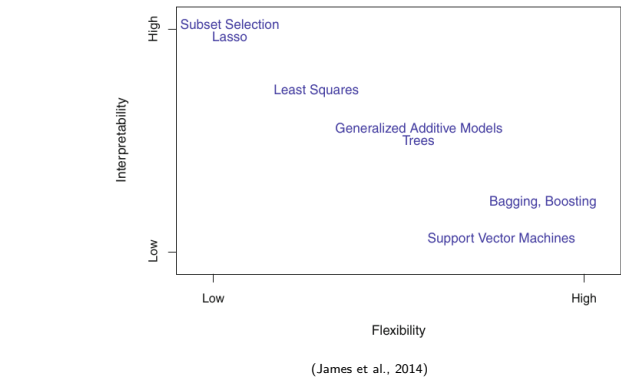
- On trouve toujours des régions d'exemples mal-classés
 - **arbre de décision** : fonction f composée de coupure perpendiculaires aux axes

Apprentissage supervisé : séparateur des classes



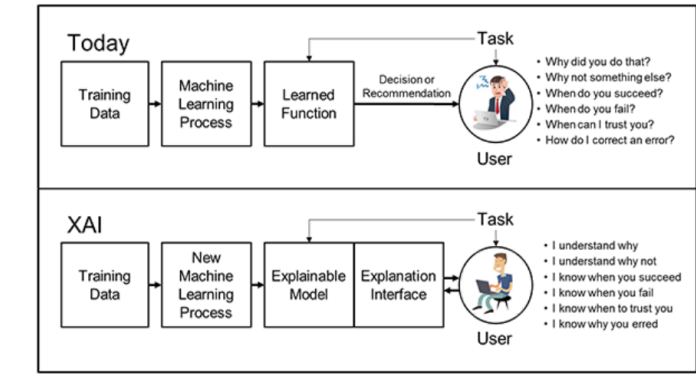
- Comment choisir un bon classifieur ?
 - taux de bonne classification (classification d'un ensemble test)
 - complexité, flexibilité
 - un objectif : **interprétabilité, explainabilité**
 - directive RGPD : règlement général sur la protection des données
<https://www.cnil.fr/fr/reglement-europeen-protection-donnees>

Apprentissage supervisé : séparateur des classes



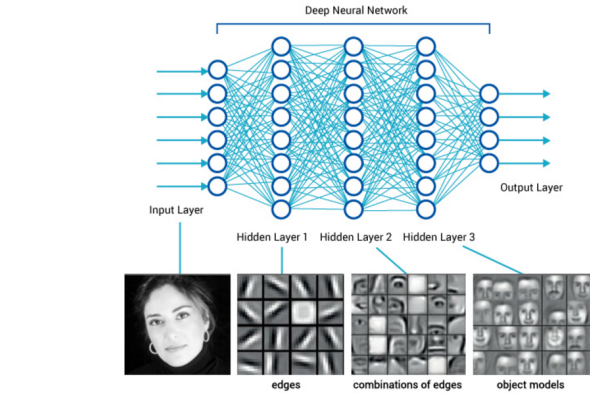
- Choix d'un classifieur : bonnes classifications ou interprétabilité ?

eXplainable AI (XAI) : Interprétabilité des systèmes



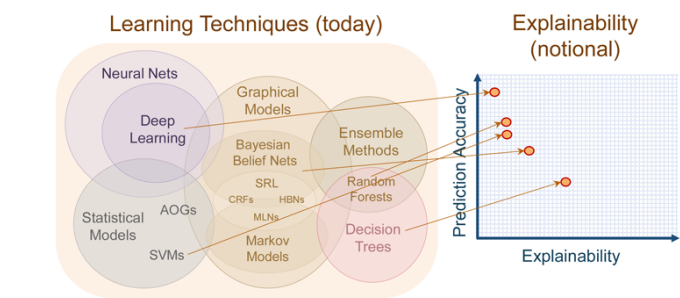
Source : (Darpa, 2016)

L'apprentissage profond : modèle général



► <https://www.amax.com>

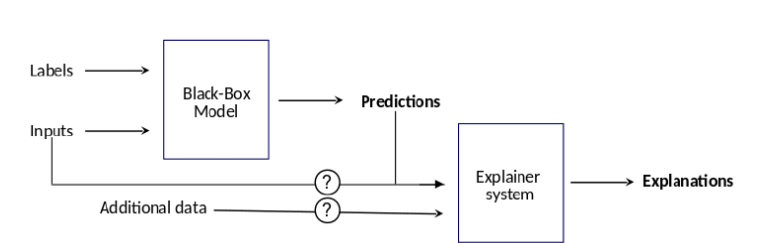
Apprentissage supervisé : séparateur des classes



<http://internetactu.blog.lemonde.fr/2016/10/30/intelligence-artificielle-va-t-elle-rester-impenetrable/>

- Choix d'un classifieur : bonnes classifications ou interprétabilité ?

Interprétabilité des systèmes : l'approche **post-hoc**

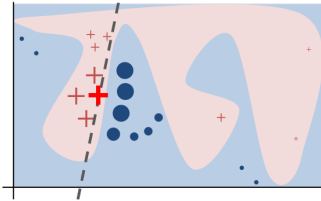


- Différentes approches :
 - "model-agnostic"
 - ou avec connaissances sur le modèle

Source : (Lauget, 2018)

Interprétabilité des systèmes : l'approche post-hoc

- Un exemple : le modèle LIME (Local Interpretable Model-agnostic Explanations)



- Estimateur linéaire local

Source : "Why should I trust you? Explaining the predictions of any classifier", Tulio Ribeiro, Singh & Guestrin, 2016

Conclusion

- Dans ce cours : introduction à l'IA et à la science des données
 - l'IA et l'apprentissage artificiel
 - problématique de l'analyse et la science des données
- Pour aller plus loin : **Master DAC** !
 - Master "Données, Apprentissage, et Connaissances"
 - IA : modélisation des raisonnements, connaissances
 - apprentissage artificiel : symbolique, numérique,...
 - bases de données, traitement, collecte, gestion de grandes quantités de données
 - recherche d'information, fouille de données
- Pour en savoir plus :
 - <http://dac.lip6.fr/master/>

Les grands défis de l'IA

- Impact sociétal de l'IA
- Interprétabilité des modèles et de leurs décisions
- IA équitable / éthique / responsable
 - respects des droits humains, de la diversité, etc.
- IA de confiance (en particulier pour des applications critiques)
 - certification, normes
 - sécurité d'utilisation
- Défis techniques
 - gestion de masses de données
 - limitation des biais d'apprentissage
 - IA "verte" et IA embarquée
 - exemple d'application : capsule endoscopique (LIP6)
<https://lejournel.cnrs.fr/articles/une-capsule-pour-cartographier-lintestin>