

L'apprentissage supervisé

Carine TOURE

LMA36-31

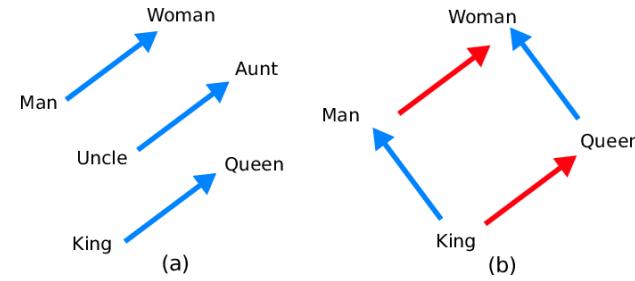
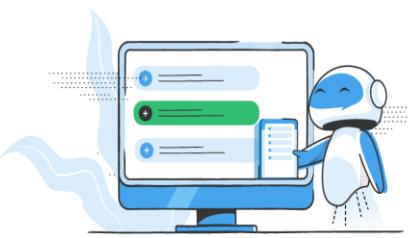
Université des Lagunes

Outline

- Préambule
 - L'apprentissage machine : vue d'ensemble
 - L'apprentissage supervisé
 - Définition
 - Exemple d'hypothèses
 - Compromis biais/variance
 - Mesure de l'erreur
 - Récapitulatif
 - Ce que nous allons voir dans la suite du cours
- Classification bayésienne
- Les k plus proches voisins
- Les arbres de décision
- La régression linéaire
- La régression logistique
- Les machines à support de vecteur

Apprentissage machine : vue d'ensemble

- Intelligence artificielle¹ => discipline générale s'intéressant à la compréhension et à la construction d'entités intelligentes (agents), ie des machines capables de calculer comment agir efficacement et de manière sûre dans une grande variété de situations nouvelles.
- Apprentissage machine (automatique) => sous-domaine de l'IA qui étudie la capacité à améliorer les performances de la machine en fonction de l'expérience. L'acquisition de compétences par l'expérience ne s'utilise pas pour toutes les méthodes.
 - IA sans apprentissage automatique : systèmes à base de règles (chatbots de 1ère génération), les plongements lexicaux (utilisés en NLP)
 - IA avec apprentissage automatique : détecteurs de spam, chatgpt, véhicules autonomes



¹ Intelligence artificielle, une approche moderne – 4ème édition 2021. Stuart Russel & Peter Novig

Apprentissage machine : vue d'ensemble

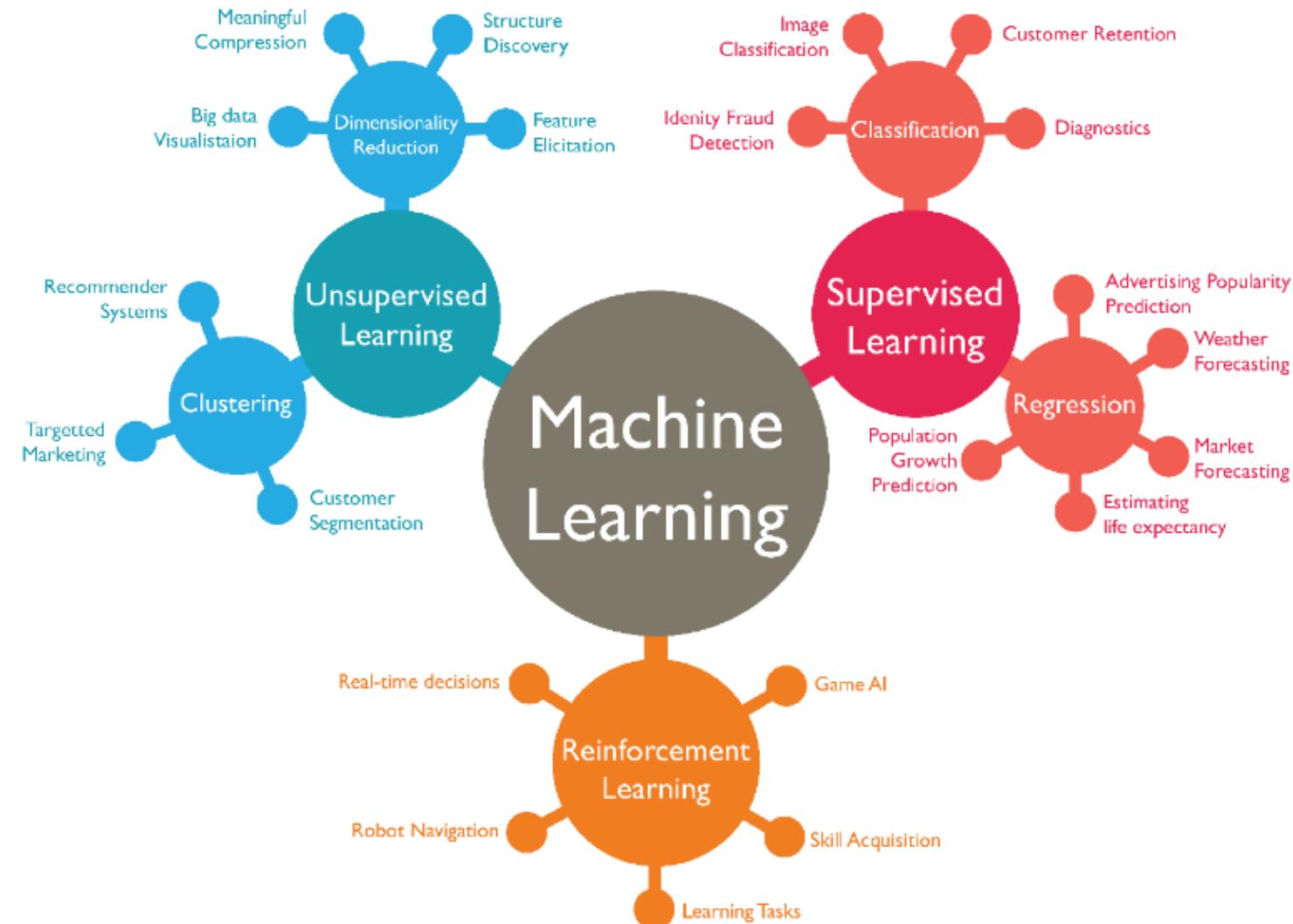


Figure 1. Les grandes familles d'apprentissage machine.
Source : <https://deeplearning.fr/tag/renforcement/>

Définition

- Apprentissage : processus d'amélioration des performances d'un agent après observation de son environnement, construction d'un modèle à partir de ces observations, utilisation de ce modèle comme hypothèse et comme procédure logicielle capable de résoudre certains problèmes.
 - Varie selon la nature de l'agent, du composant à améliorer et du retour disponible
- Si le retour disponible indique la réponse correcte pour des exemples d'entrée, on parle d'**apprentissage supervisé**.
- Pour un ensemble d'apprentissage de N exemples couples entrée-sortie $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$:
 - y_i : référence, réponse véritable que l'on cherche à prédire
 - Chaque y_i a été produit par une fonction inconnue $y = f(x)$
 - Objectif : trouver une fonction h (l'hypothèse cohérente) qui approche la vraie fonction f . h provient d'un espace d'hypothèses H (ou classe de fonctions) qui contient les fonctions possibles.
- Lorsque les valeurs de y sont discrètes, on a un problème de **classification**

Intelligence artificielle, une approche moderne - 4ème édition 2021 - Stuart Russell et Peter Norvig

Définition

■ Comment choisir l'espace d'hypothèses ?

- connaissance a-priori du processus ayant engendré les données, exploration par analyse des données (visualisations, tests statistiques, ...)

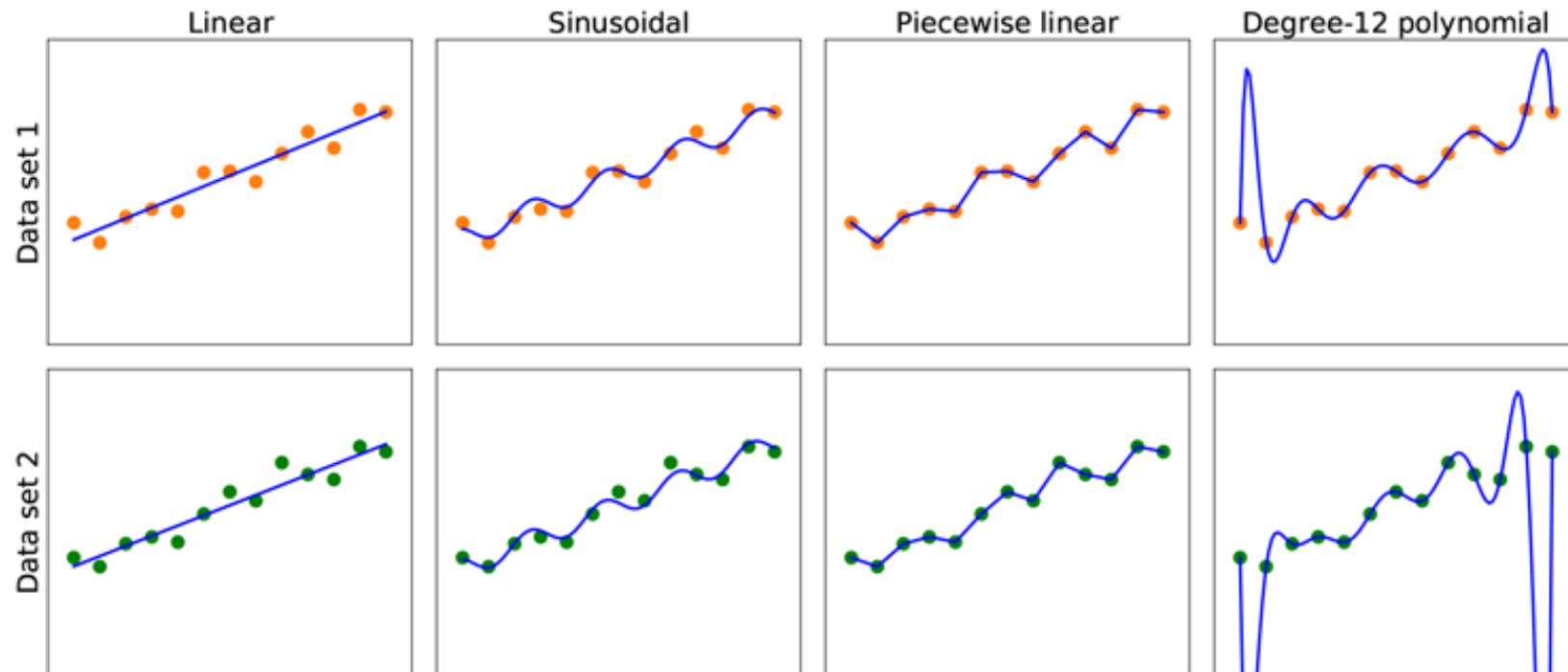


Figure 2. Exemples d'hypothèses pour approcher les données.

Rangée du haut : quatre tracés avec la meilleure hypothèse issue d'un des quatre espaces d'hypothèses, par entraînement sur un ensemble de données

Rangée du bas : la même chose, mais avec un échantillon différent du même ensemble de données

Définition

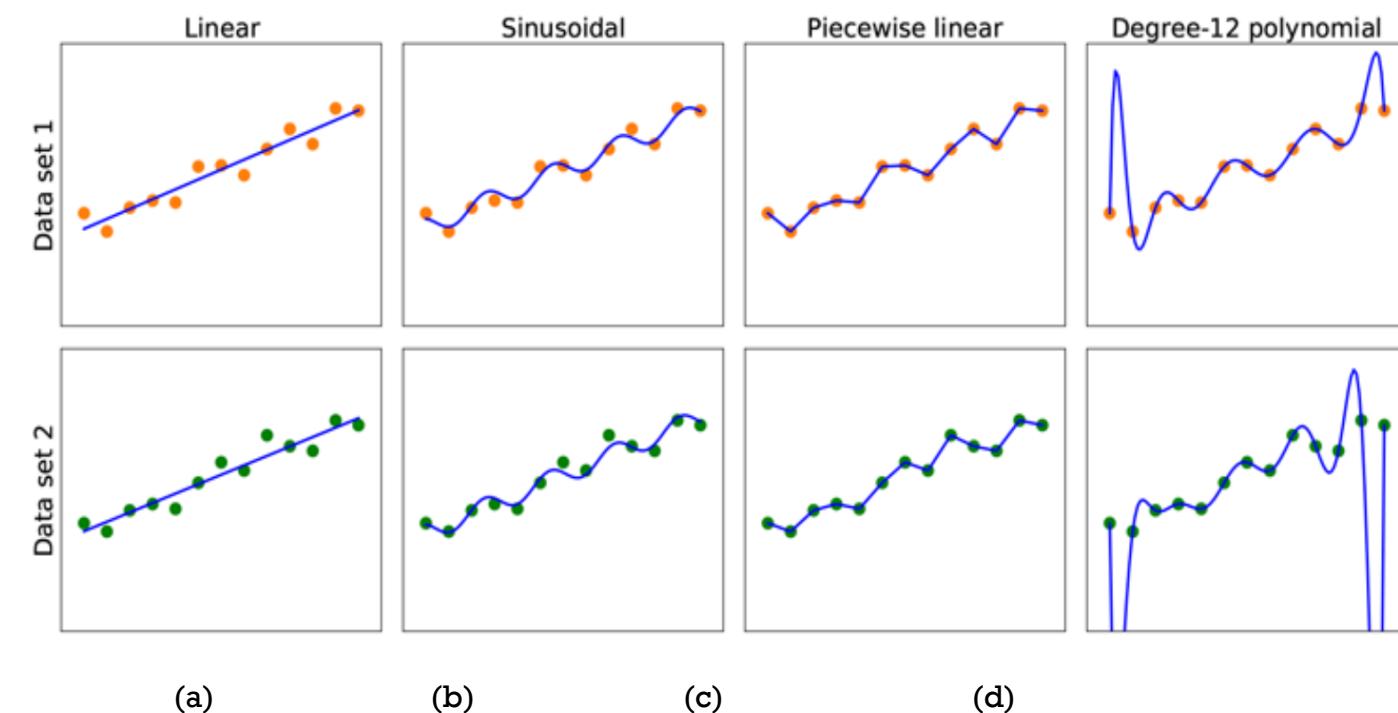
Exemple d'hypothèses

Colonne (a) : Lignes droites; fonctions de la forme $h(x) = w_1x + w_0$. Aucune ligne droite n'est une hypothèse cohérente pour ces données

Colonne (b) : Sinusoïdes; fonctions de la forme $h(x) = w_1x + \sin(w_0x)$. Ce choix n'est pas parfaitement cohérent mais s'adapte bien aux données.

Colonne (c) : Fonctions linéaires par morceaux; chaque segment relie simplement un point à un autre. Ces fonctions sont toujours cohérentes par nature.

Colonne (d) : Polynômes de degré 12; $h(x) = \sum_{i=0}^{12} w_i x^i$. Ils sont cohérents puisqu'on peut toujours faire passer exactement un polynôme de degré 12 par 13 points distincts.



Attention ! La cohérence de l'hypothèse s'assure pas sa qualité

Nécessité d'analyser le biais et la variance qu'ils imposent

Compromis biais/variance

Le biais

La variance

Compromis biais/variance :

La mesure de l'erreur

- Comment choisir l'hypothèse cohérente dans l'espace d'hypothèses ?
 - Trouver h telle que pour chaque x de l'ensemble d'entraînement, on a $h(x_i) \approx y_i$
 - Pour cela, minimiser une fonction de perte $L(x, y, \hat{y})$
- La mesure objective de l'hypothèse ne se fait pas sur son comportement sur l'ensemble d'entraînement mais sur un ensemble de test.
- On dit que h généralise correctement si elle prédit correctement les sorties sur l'ensemble de test

Pour récapituler

- L'ensemble des observations (dataset)
- L'ensemble d'entraînement (training set)
- L'ensemble de test (test set)
- La fonction hypothèse ()
- La fonction de perte (loss function)
- Le biais (bias)
- La variance (variance)
- Sur-adaptation (overfitting)
- Sous-adaptation (underfitting)
- Le compromis biais-variance (bias/variance trade-off)
- La régression
- La classification

¹ Intelligence artificielle, une approche moderne – 4ème édition 2021. Stuart Russel & Peter Novig

Ce que nous allons voir

- CM / TD / TP
- Rapporteurs
- Objectifs
- Évaluation

Outline

- Préambule
- Classification bayésienne
 - Théorème de Bayes
 - Performance du modèle
 - Optimisation du modèle
- Les k plus proches voisins
- Les arbres de décision
- La régression linéaire
- La régression logistique
- Les machines à support de vecteur