

Régression linéaire

Noura BENHAJJI

noura.benhajji@gmail.com

Introduction

- ❑ La régression sert à trouver la **relation** d'une variable par rapport à une ou plusieurs autres.
- ❑ Le but de la régression est d'estimer une **valeur** (numérique) de sortie à partir des valeurs d'un ensemble de caractéristiques en entrée.
- ❑ Le problème revient à estimer une fonction de calcul en se basant sur les données d'entraînement.

Algorithmes de régression

❑ Régression linéaire

- ❑ Trouver une droite au milieu d'un nuage de points.

❑ Régression polynomiale

- ❑ Trouver une relation de forme non-linéaire entre la réponse (y) et la ou les variables explicatives (x)
- ❑ pour prendre en charge cette forme non linéaire, les algorithmes de régressions intègrent des polynômes dans leurs équations.

Algorithmes de classification supervisée

❑ Le random Forest

- ❑ calculer la moyenne des prévisions obtenues

❑ Le perceptron multicouches

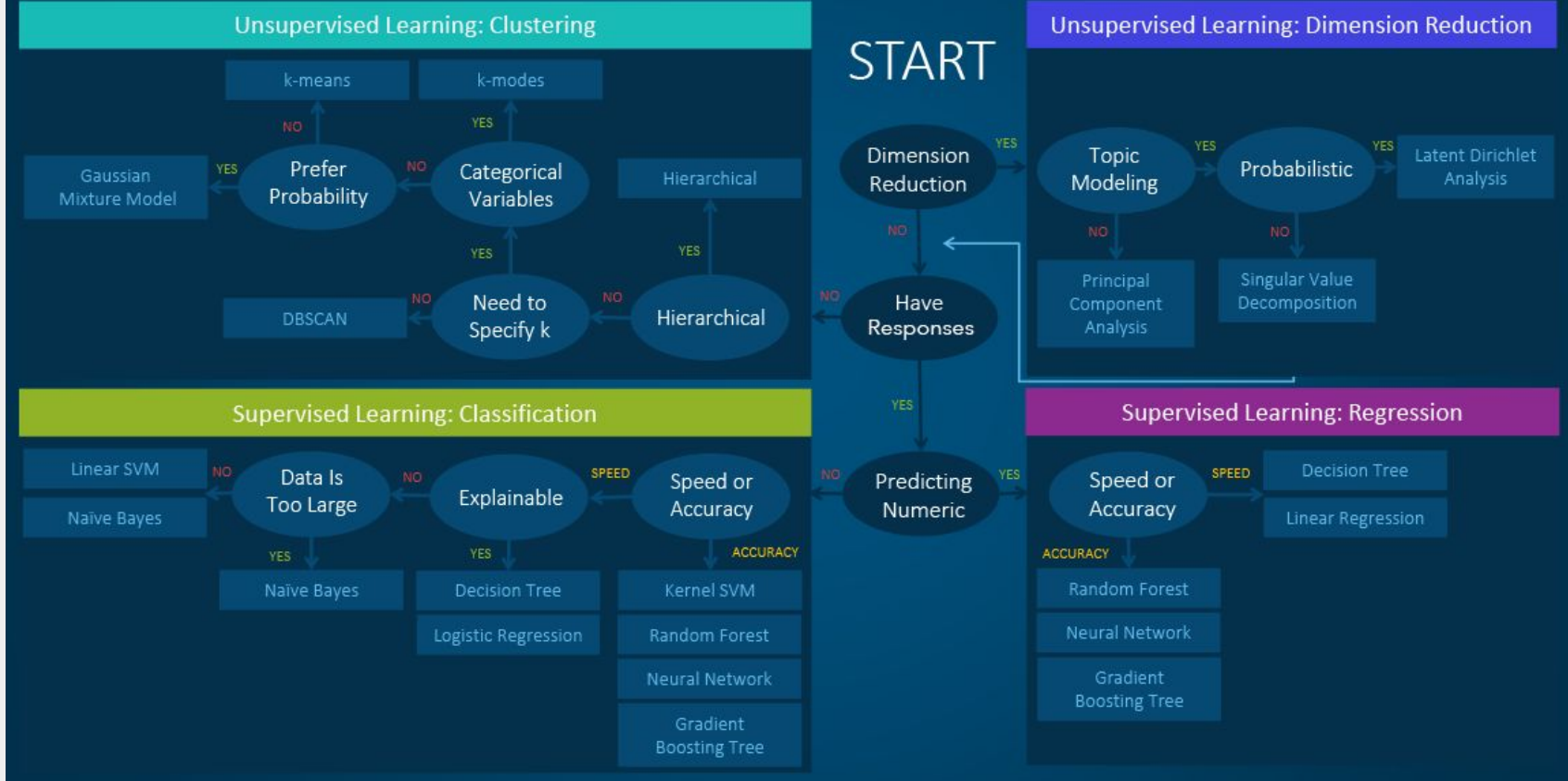
- ❑ Un ensemble de neurones connectés.
- ❑ Il est composé d'une couche d'entrée, de n couches cachées, et d'une couche de sortie.

Algorithmes de classification supervisée

❑ SVM

- ❑ le principe des SVM consiste à ramener un problème de classification à un hyperplan dans lequel les données sont séparées en plusieurs classes dont la frontière est la plus éloignée possible des points de données.
- ❑ un SVM permet de résoudre les problèmes de régression en utilisant la technique du noyau (à venir)

Machine Learning Algorithms Cheat Sheet



Source : SAS Algorithm Flowchart



Régression linéaire

- ❑ Qu'est ce qu'une régression linéaire ?
- ❑ Quelles sont les étapes de construction d'un modèle régression linéaire ?
 - ❑ fonction coût
 - ❑ minimiser la fonction coût
 - ❑ méthode de résolution du problème (descente de gradient, méthode des moindres carrés)
- ❑ préparation des données

Ressources :

- [La régression linéaire pour comprendre les grands principes du Machine Learning](#)
- [Comprendre la descente de gradient en 3 étapes et 12 dessins](#)
- [L'algorithme de Gradient Descent](#)
- [Notions fondamentales de la régression linéaire](#)
- [Préparation des données](#)



Travail en groupe



45min

Régression linéaire

- ❑ Notebook : 1. Régression linéaire multiple
- ❑ Veille : matrice de corrélation



Veille individuelle



45 min

Comment évaluer un modèle de régression ?

- ❑ Coefficient de détermination
- ❑ Mean Absolute Error
- ❑ Mean Squared Error
- ❑ Root Mean Square Error



Travail en groupe



Evaluation d'un modèle de régression

- ❏ Notebook : [Régression linéaire] Métriques - Apprenants



Veille individuelle



60 min

Sélection de variables

- ❑ Quel intérêt ?
- ❑ Filter Method
- ❑ Wrapper Method
- ❑ Embedded Method



Travail en groupe



60 min

Sélection de variables

- ❏ Tuto : Feature selection
python datacamp : A Case
study in Python



Veille individuelle



60 min

Hypothèses de régression linéaire

- ❑ Exogénéité
- ❑ Homoscédasticité
- ❑ Erreurs indépendantes
- ❑ Normalité des erreurs
- ❑ Non colinéarité des variables indépendantes



Travail en groupe



Modèle de prédiction de prix de diamants

- ❑ data
- ❑ Etapes :
 - ❑ exploration des données
 - ❑ préparation de données
 - ❑ modèle 1 : régression linéaire sans sélection de variables
 - ❑ évaluation modèle 1
 - ❑ interprétation modèle 1 (coef)
 - ❑ modèle 2 : régression linéaire avec sélection de variables
 - ❑ évaluation modèle 2
 - ❑ comparer modèle 1 et modèle 2
 - ❑ interprétation modèle 1 (coef)

