

# Machine Learning en assurance vie : application à la valeur client

---



**Candice Elmaleh** (Abeille Assurances)

**Ludovic Moreau** (Abeille Assurances)

**Charles Boddèle** (Finactys)

**Damien Loureiro** (Finactys)



**Introduction**

1. Valeur client
2. Clustering
3. Prédiction des flux futurs

Conclusions

# Introduction

---

## Introduction

1. Valeur client
2. Clustering
3. Prédiction des flux futurs

Conclusions



**Des volumes de données** de plus en plus conséquents

### **Non-vie**

Événements fréquents (sinistres, réclamations, etc.)

### **Vie**

Événement moins fréquents (décès, rachats, etc.) mais variables complémentaires : profil investisseur, etc.



Exploités par du **Machine Learning**

### **Apprentissage automatique**

Approche basée sur les données  
Analyse les données pour déceler des patterns

### **Des modèles variés**

Des modèles adaptés aux besoins et aux données : apprentissage supervisé/non supervisé



Pour une **application concrète** en assurance

### **Non-vie**

Tarification dynamique  
Détection de fraude  
Prédiction de sinistre

### **Vie**

Application à la valeur client

Introduction

**1. Valeur client**

2. Clustering

3. Prédiction des  
flux futurs

Conclusions

# Valeur client

---

Introduction

**1. Valeur client**

2. Clustering

3. Prédiction des flux futurs

Conclusions

Portefeuille  
**clients**



# Valeur client | Définition

Introduction

## 1. Valeur client

2. Clustering

3. Prédiction des flux futurs

Conclusions

Portefeuille  
clients



Valeur client  
*contribution financière*

$V_1$

$V_2$

$V_3$

$V_N$



Introduction

## 1. Valeur client

2. Clustering

3. Prédiction des flux futurs

Conclusions

Portefeuille  
clients



Valeur client  
*contribution financière*

$V_1$

$V_2$

$V_3$

$V_N$



Objectifs  
principaux

1

2

3

# Valeur client | Définition

Introduction

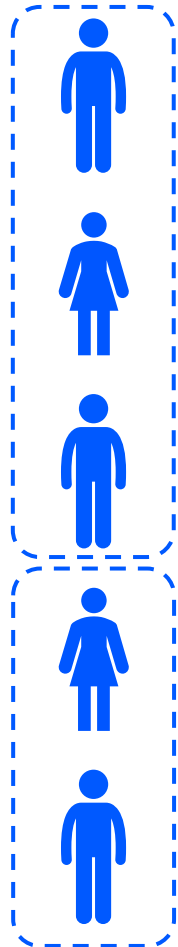
## 1. Valeur client

2. Clustering

3. Prédiction des flux futurs

Conclusions

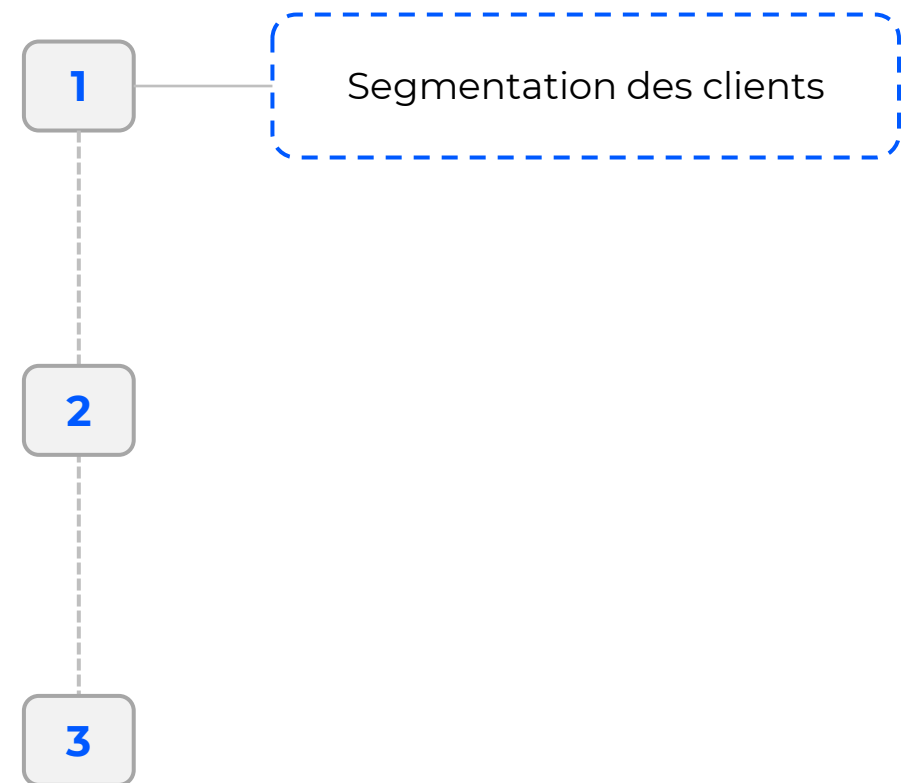
### Portefeuille clients



### Valeur client *contribution financière*



### Objectifs principaux





Introduction

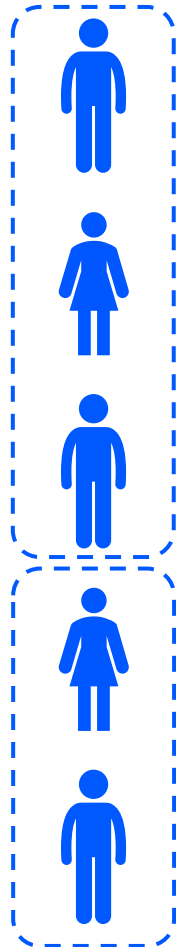
## 1. Valeur client

2. Clustering

3. Prédiction des flux futurs

Conclusions

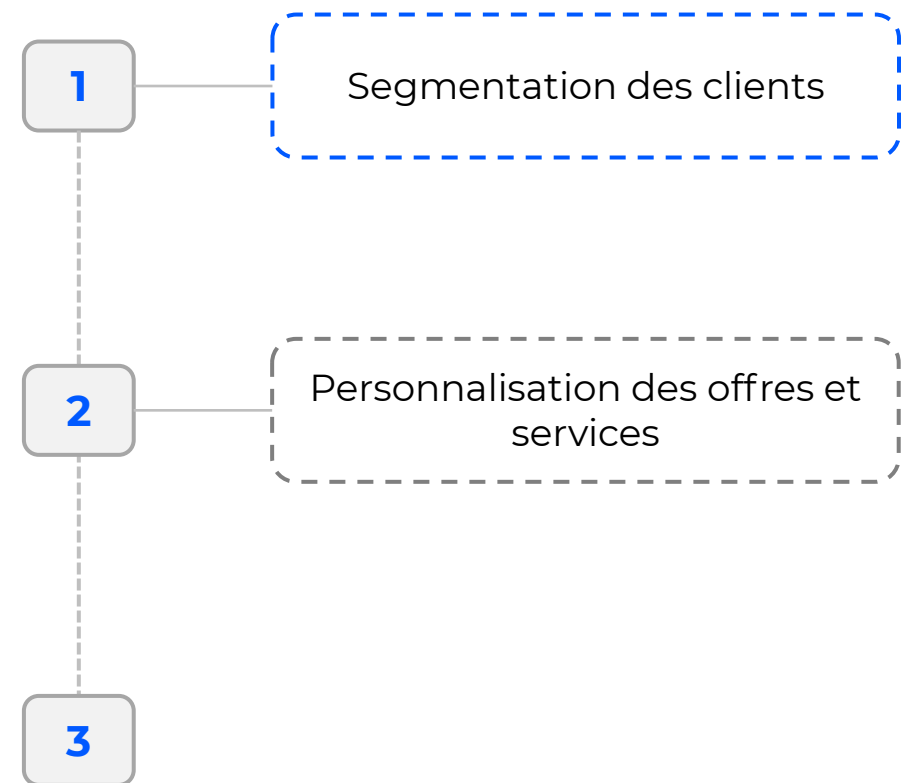
### Portefeuille clients



### Valeur client *contribution financière*



### Objectifs principaux



Introduction

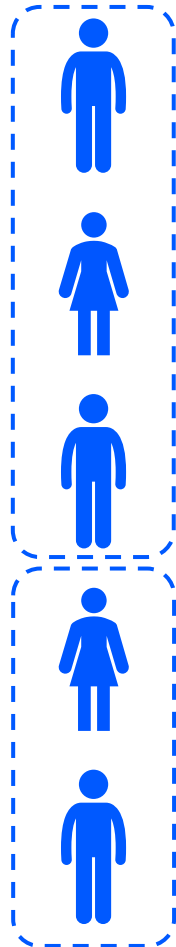
## 1. Valeur client

2. Clustering

3. Prédiction des flux futurs

Conclusions

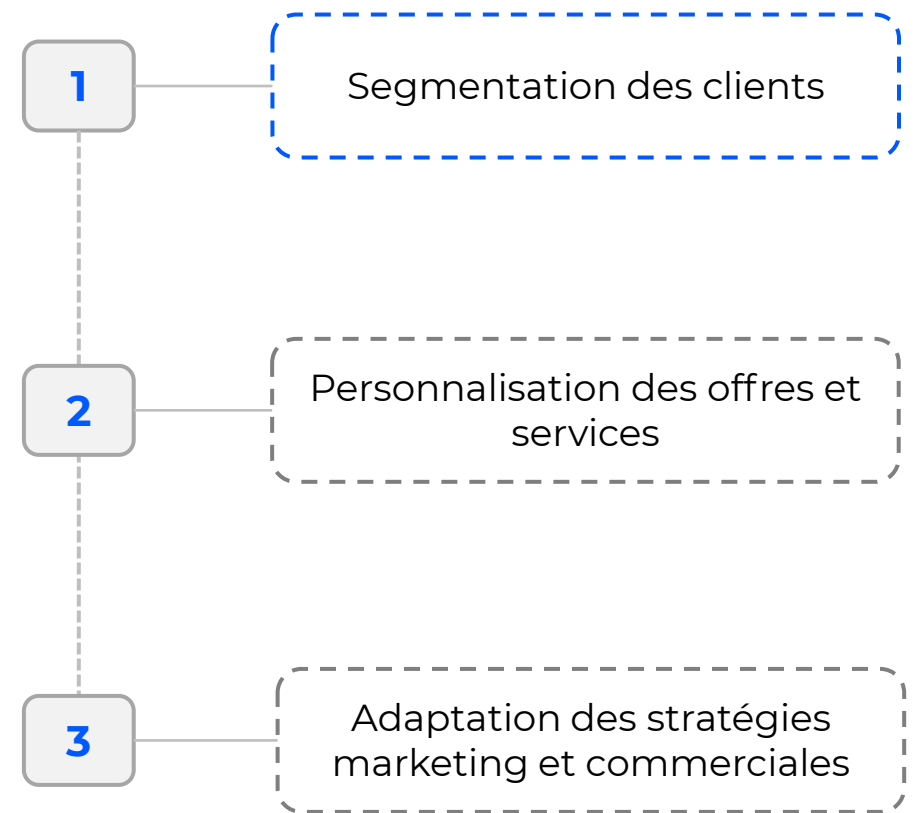
### Portefeuille clients



### Valeur client *contribution financière*



### Objectifs principaux



Introduction

## 1. Valeur client

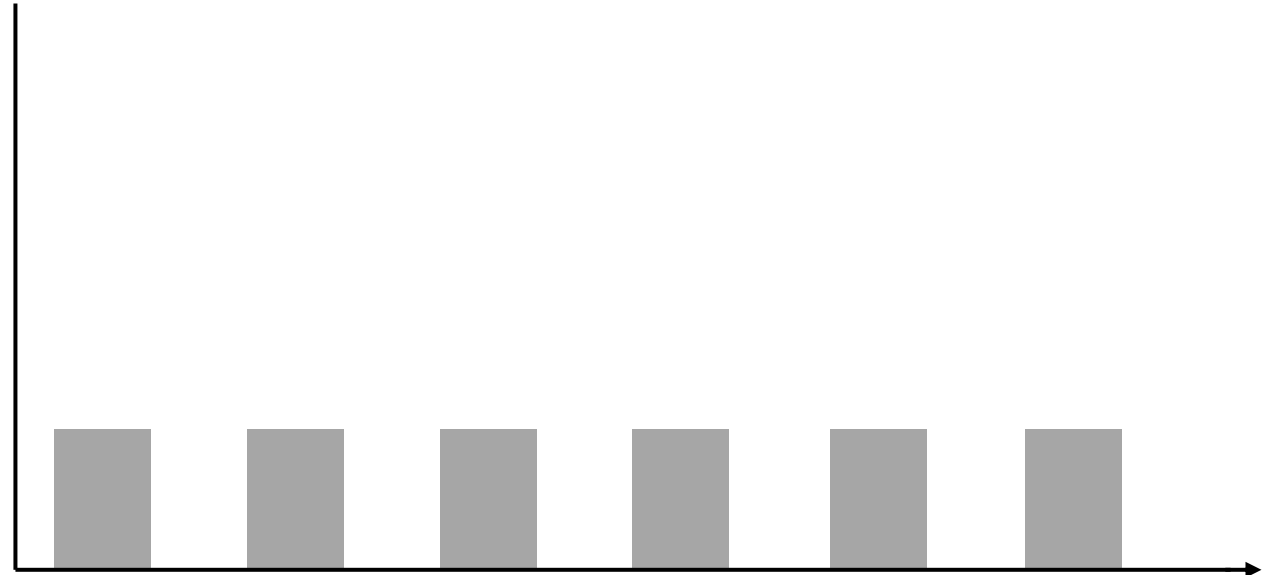
2. Clustering

3. Prédiction des flux futurs

Conclusions

### Valeur client

- Revenus futurs attendus



### Contrat en stock

Contrat d'assurance vie

Calcul des résultats futurs

Introduction

## 1. Valeur client

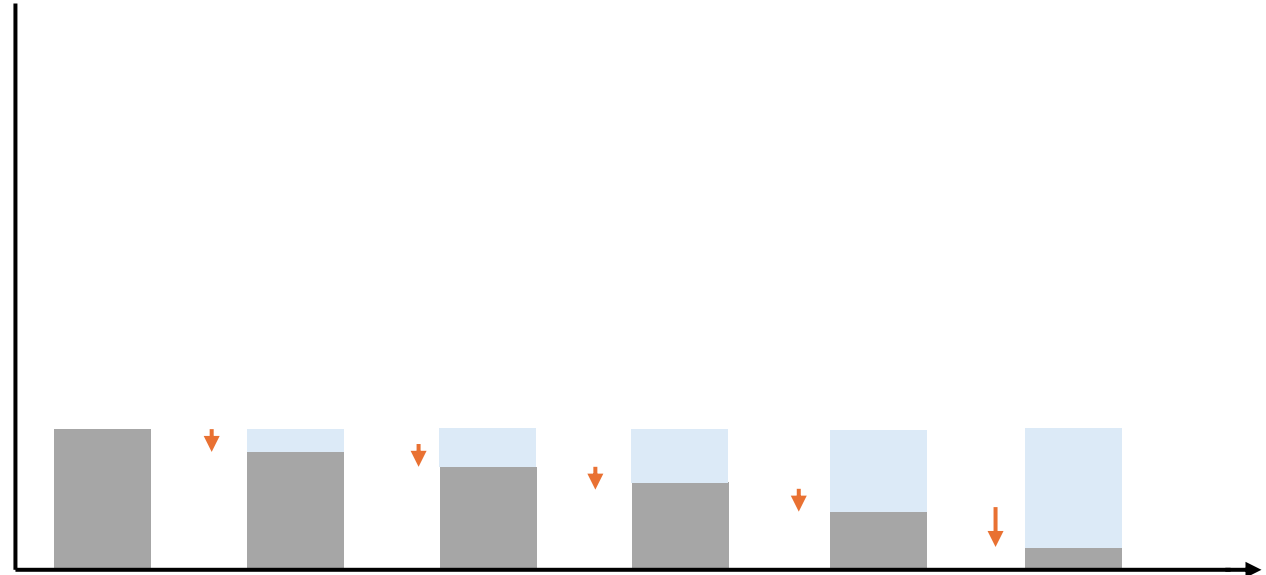
2. Clustering

3. Prédiction des flux futurs

Conclusions

### Valeur client

- Revenus futurs attendus
- Fidélité de client



### Contrat en stock

Contrat d'assurance vie  
Calcul des résultats futurs

### Rachats / décès

Montant bien connu : VIF  
Calcul à la maille client ⚙️

Introduction

## 1. Valeur client

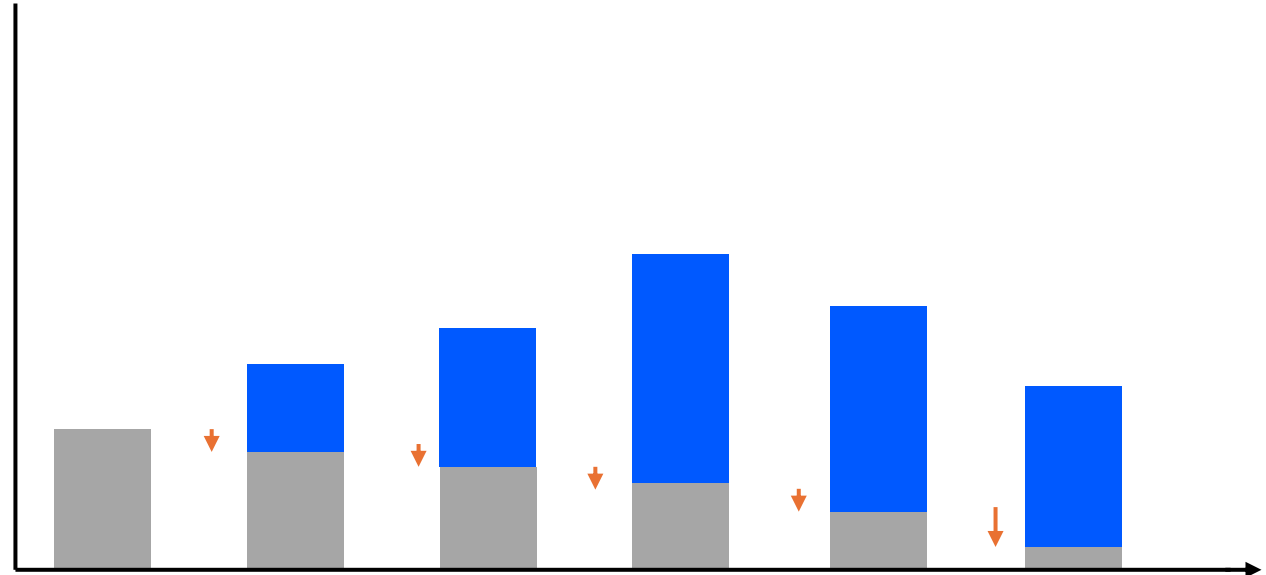
2. Clustering

3. Prédiction des flux futurs

Conclusions

### Valeur client

- Revenus futurs attendus
- Fidélité de client
- Potentiel de **croissance**



#### Contrat en stock

Contrat d'assurance vie  
Calcul des résultats futurs

#### Rachats / décès

Montant bien connu : VIF  
Calcul à la maille client ⚙️

#### Primes futures

Versements ⚙️  
Cross sell ⚙️

Introduction

1. Valeur client

**2. Clustering**

3. Prédiction des  
flux futurs

Conclusions

# Clustering

---

Introduction

1. Valeur client

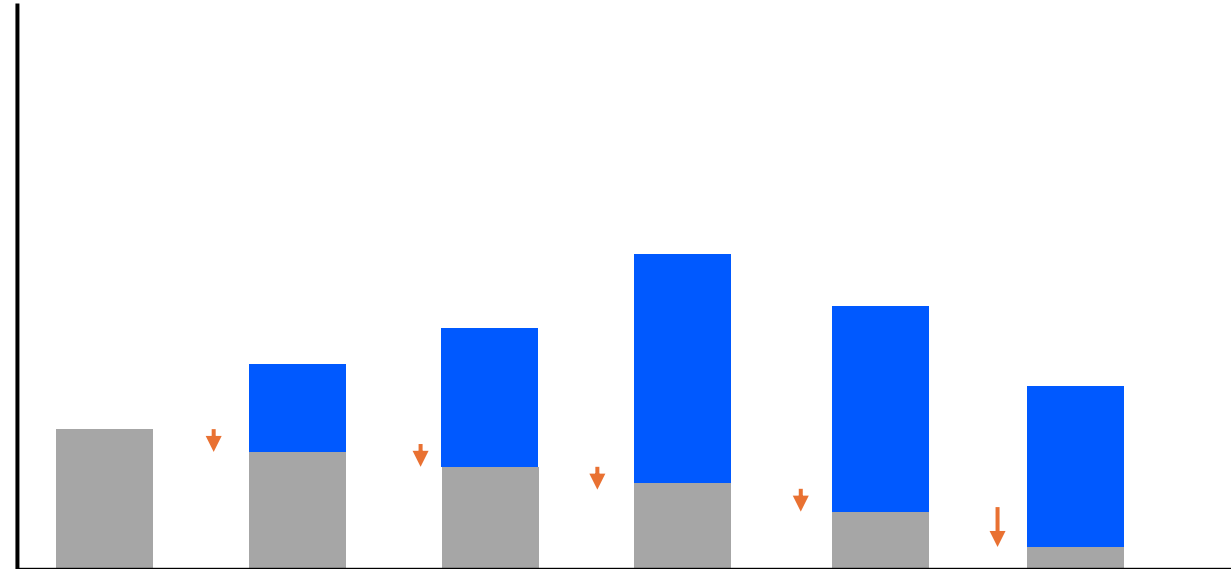
**2. Clustering**

3. Prédiction des flux futurs

Conclusions

## Valeur client

- **Revenus futurs** attendus
- **Fidélité** de client
- Potentiel de **croissance**



Introduction

1. Valeur client

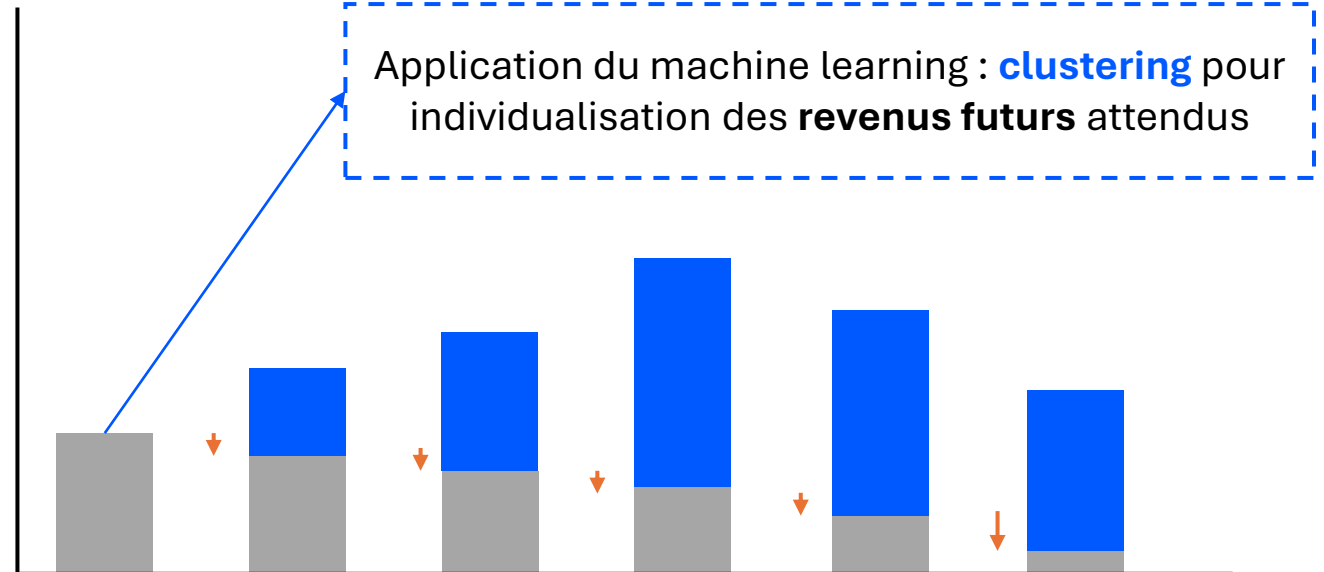
**2. Clustering**

3. Prédiction des flux futurs

Conclusions

## Valeur client

- **Revenus futurs** attendus
- **Fidélité** de client
- Potentiel de **croissance**



Utilisation de modèles non supervisés



# Clustering | groupe homogène de clients

Application du machine learning : **clustering** pour individualisation des **revenus futurs** attendus

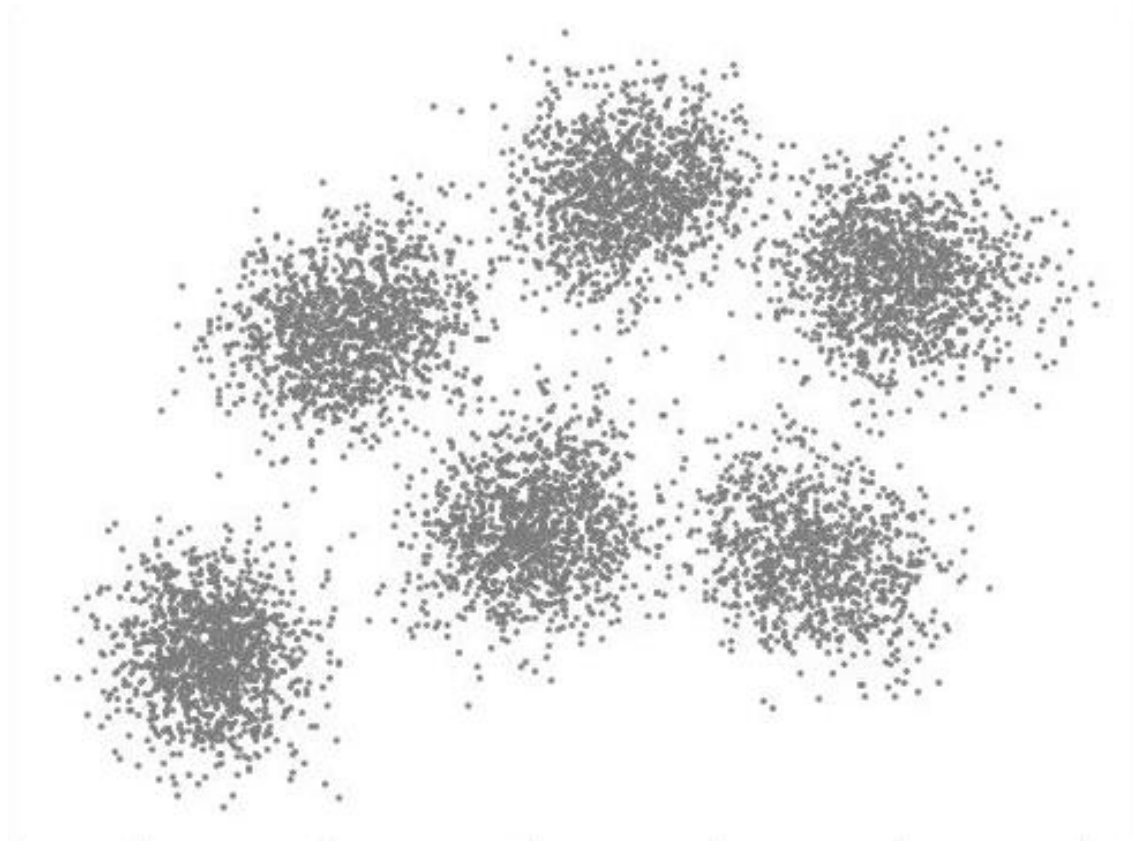
Introduction

1. Valeur client

**2. Clustering**

3. Prédiction des flux futurs

Conclusions



**Client par client**

Population de X clients

Lourdeur de calcul

**Clustering pour regrouper** les clients qui se ressemblent

Introduction

1. Valeur client

**2. Clustering**

3. Prédiction des flux futurs

Conclusions

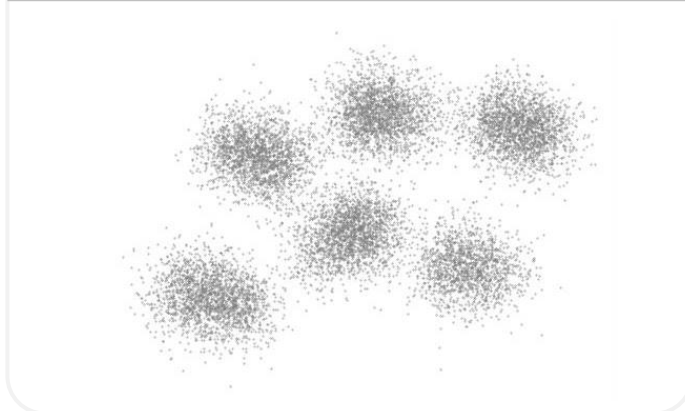
**Objectif** : identifier et regrouper des individus similaires selon plusieurs critères

Généralement, deux individus sont similaires lorsqu'ils sont proches au sens de la distance. L'algorithme cherche donc à :

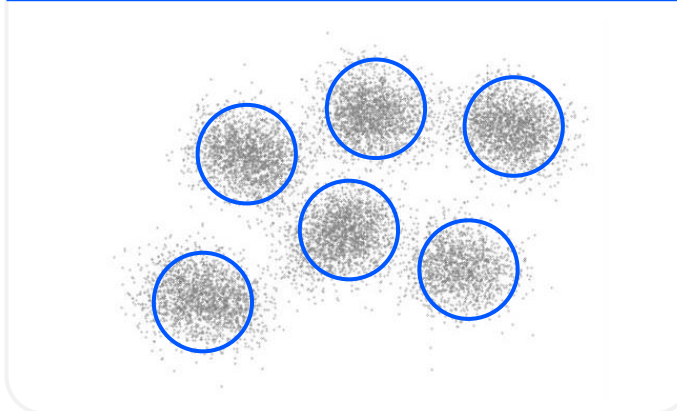
- Minimiser la **variance intra-classe** > les individus d'un même groupe doivent être le plus homogène que possible
- Maximiser la **variance inter-classe** > les groupes construits doivent être les plus séparés que possible

**Avantages par rapport à une classification humaine** : capacité à traiter beaucoup de variables et à avoir une vision large de l'ensemble des variables, éviter le biais humain, gain de temps, possibilité de fixer un nombre maximum de clusters, ...

Données à l'origine



Bon clustering



Mauvais clustering



# Clustering | groupe homogène de clients

Application du machine learning : **clustering** pour individualisation des **revenus futurs** attendus

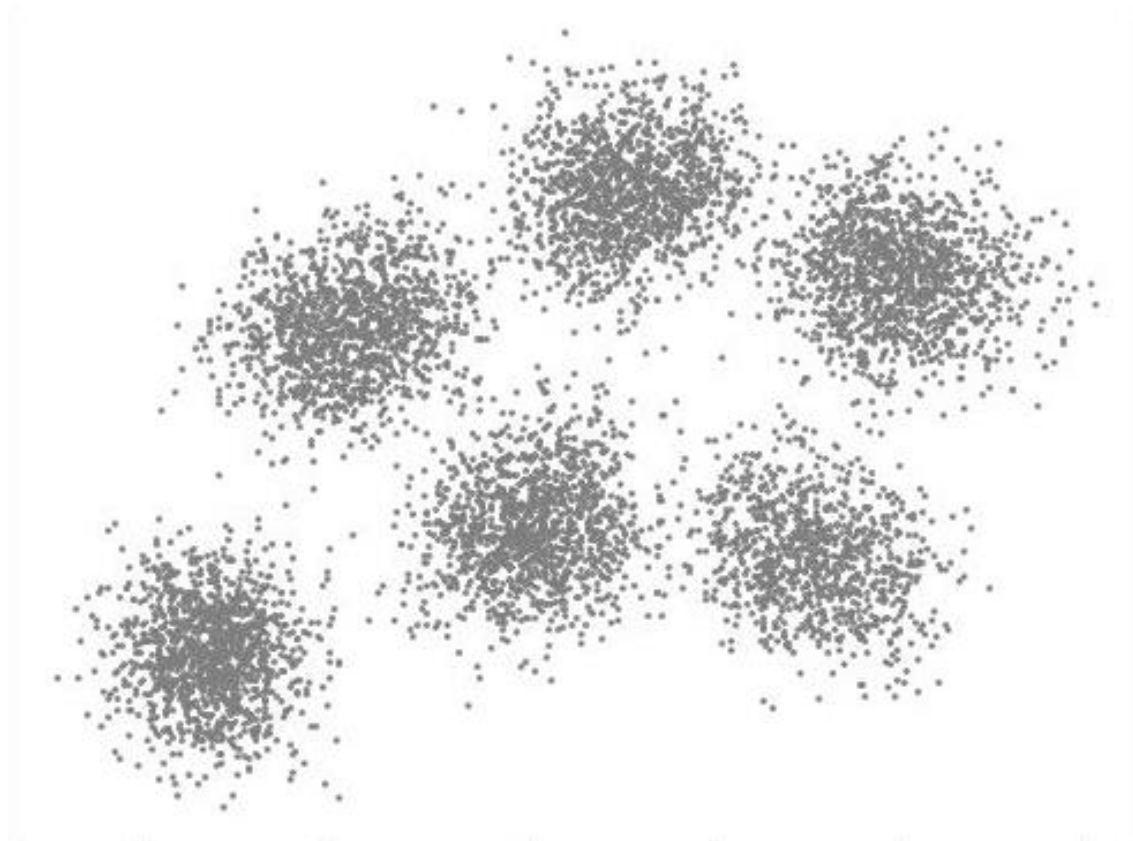
Introduction

1. Valeur client

**2. Clustering**

3. Prédiction des flux futurs

Conclusions



**Client par client**

**Critères retenus** pour le clustering

Typologie de clients

Produits et tarification

Comportement passé et prédit

# Clustering | groupe homogène de clients

Application du machine learning : **clustering** pour individualisation des **revenus futurs** attendus

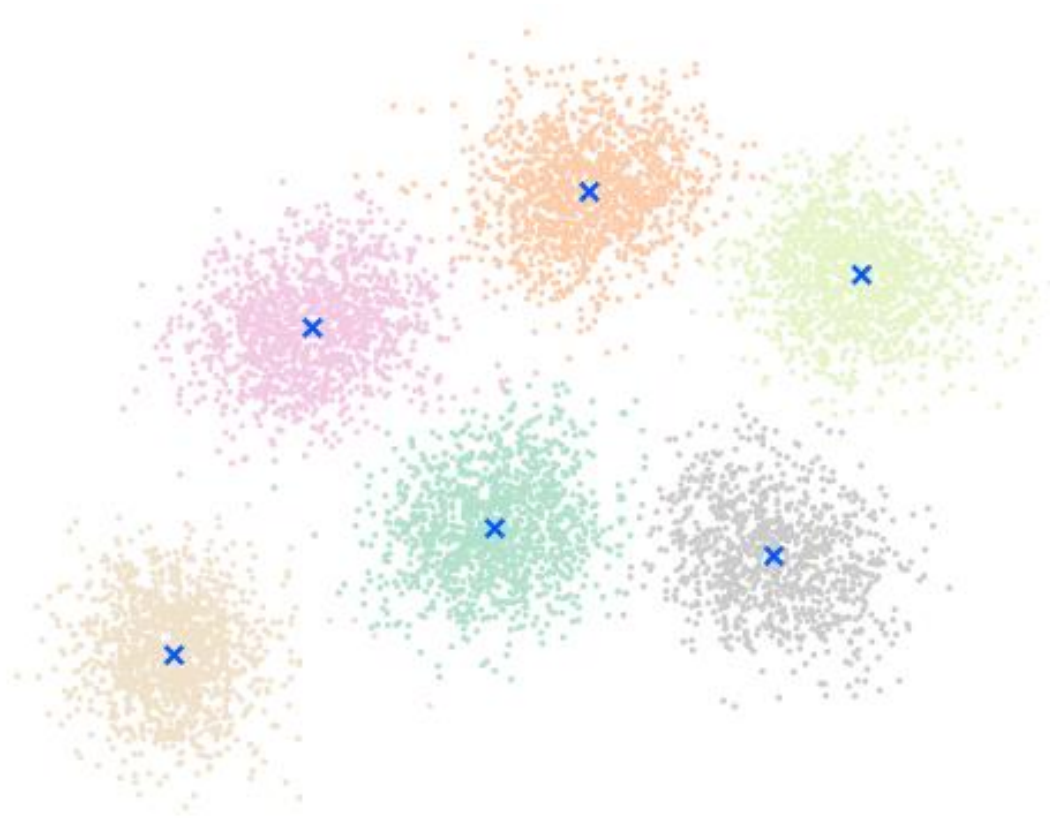
Introduction

1. Valeur client

**2. Clustering**

3. Prédiction des flux futurs

Conclusions



Groupes de clients homogènes

**Critères retenus** pour le clustering

Typologie de clients

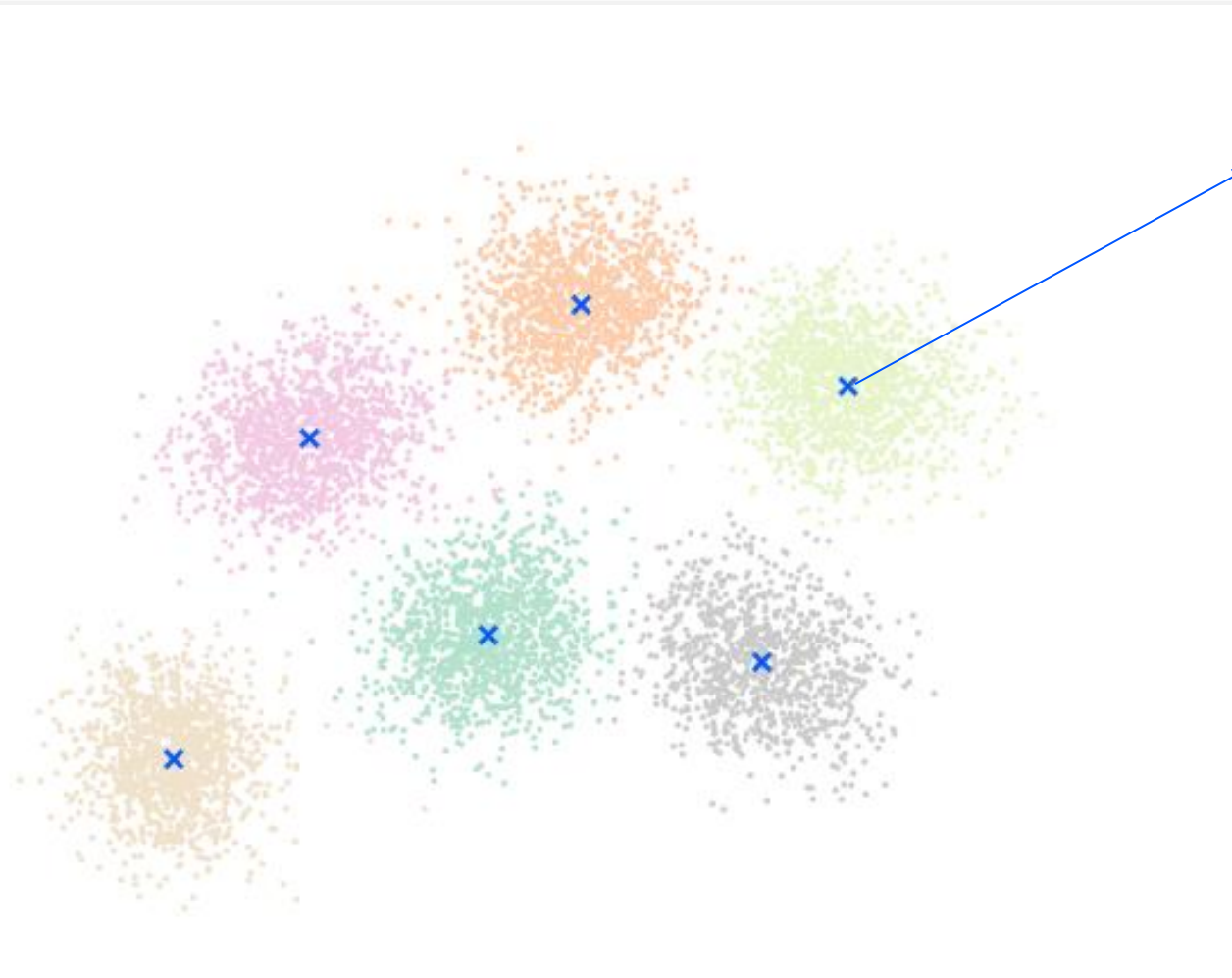
Produits et tarification

Comportement passé et prédit

Chaque groupe homogène à la **même valeur client**

Application du machine learning : **clustering** pour individualisation des **revenus futurs** attendus

- Introduction
- 1. Valeur client
- 2. Clustering**
- 3. Prédiction des flux futurs
- Conclusions

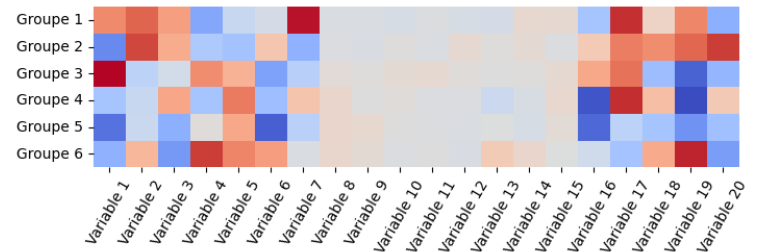


## Analyse des centroïdes pour comparaison des clusters

Quel cluster a la meilleure valeur client ?

Quelles variables ont impactées la construction de ce cluster ?

Exemple d'analyse avec un corrélogramme



Vérification et validation par la connaissance métier / business

Groupes de clients homogènes

Introduction

1. Valeur client

2. Clustering

**3. Prédiction des  
flux futurs**

Conclusions

# Prédiction des flux futurs

---

Introduction

1. Valeur client

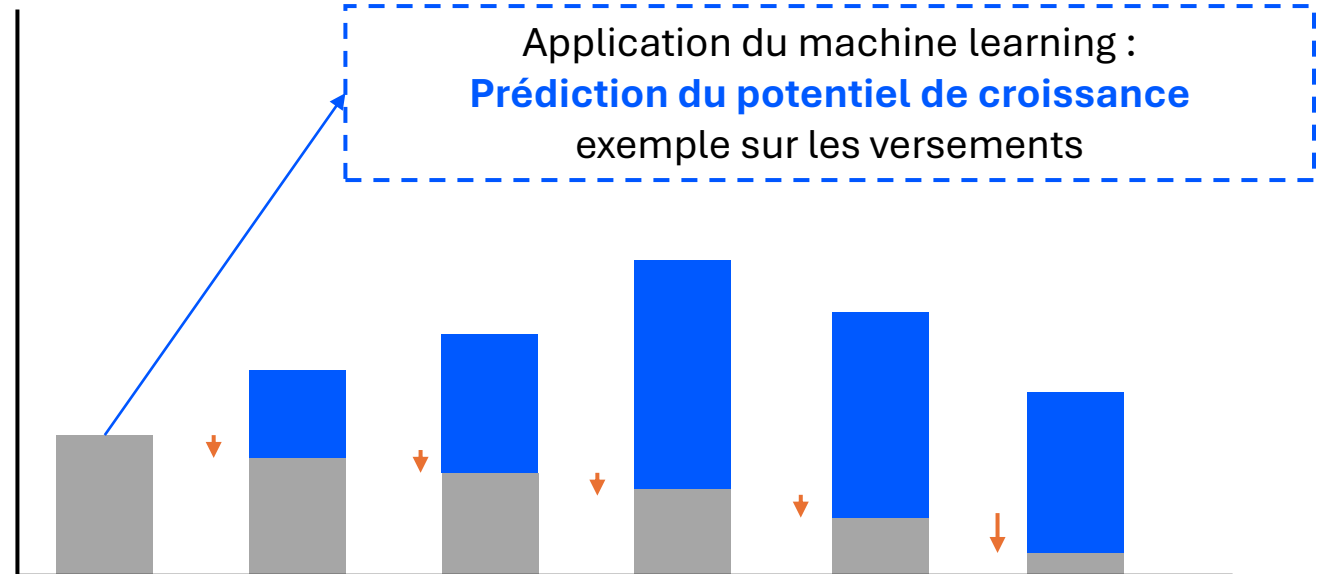
2. Clustering

**3. Prédiction des flux futurs**

Conclusions

## Valeur client

- Revenus futurs attendus
- Fidélité de client
- Potentiel de **croissance**



Utilisation de modèles supervisés



Application du machine learning :  
**Prédiction du potentiel de croissance**, exemple  
sur les versements

Introduction

1. Valeur client

2. Clustering

**3. Prédiction des flux futurs**

Conclusions

## GLM

Les GLM sont une généralisation de la régression linéaire en permettant notamment de transformer les données pour que les prédictions respectent les contraintes des réponses.

- + Facilement explicable grâce aux coefficients
- Généralement peu performant
- Beaucoup d'hypothèses sont nécessaires

## Random Forest

Le Random Forest est un algorithme de Machine Learning qui agrège plusieurs arbres de décision pour améliorer la précision et la robustesse des prédictions.

- + Généralement performant
- + Pas d'hypothèse paramétrique
- Difficilement explicable

## XGBoost

XGBoost est un algorithme de ML qui combine plusieurs arbres de décision, chaque arbre tentant de corriger les erreurs des précédents.

- + Généralement très performant
- + Temps d'exécution rapide
- Difficilement explicable
- Plusieurs hyperparamètres à optimiser

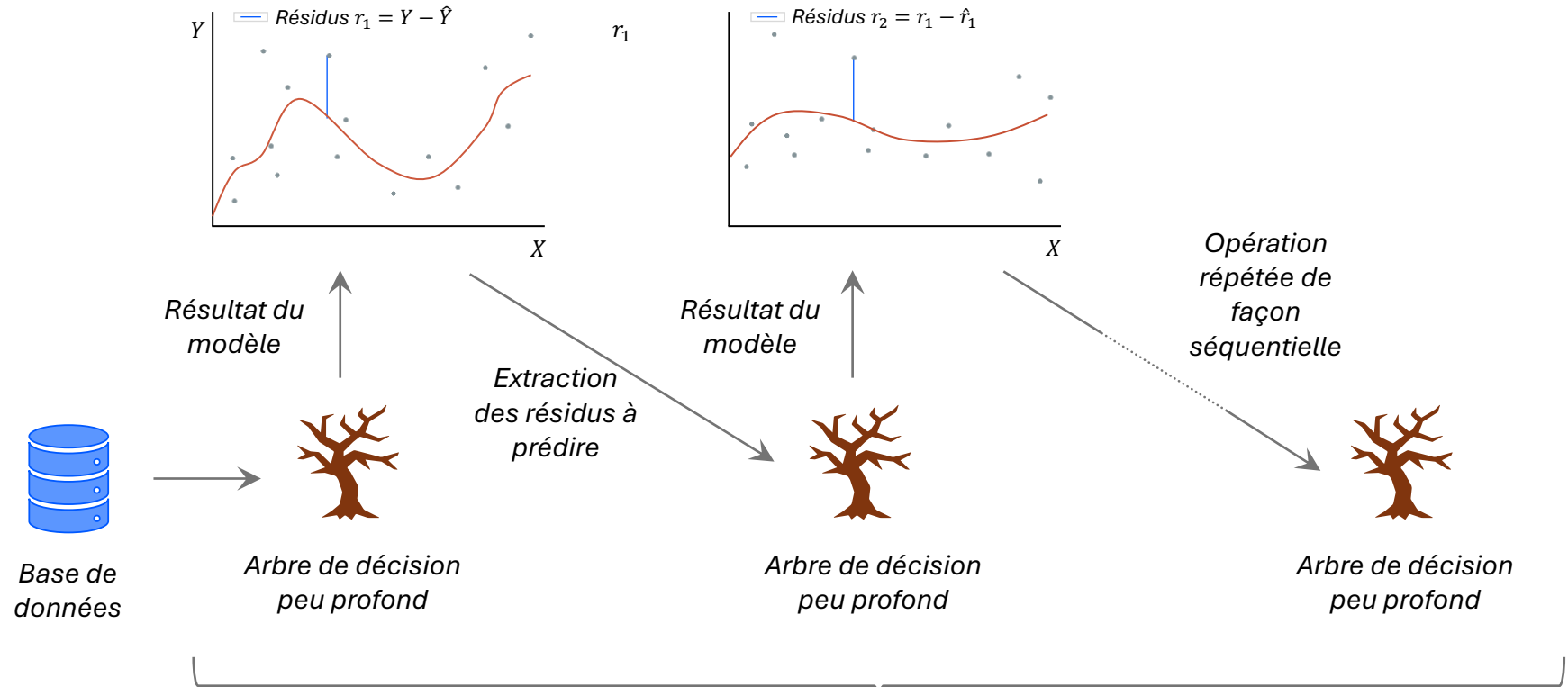
Modèle retenu



# Flux futurs | Principe du Modèle XGBoost

Application du machine learning :  
**Prédiction du potentiel de croissance**, exemple  
sur les versements

- Introduction
- 1. Valeur client
- 2. Clustering
- 3. Prédiction des flux futurs**
- Conclusions



La prédiction finale est une combinaison linéaire des modèles construits, avec des coefficients optimisés

Limites


Introduction

1. Valeur client

2. Clustering

**3. Prédiction des flux futurs**

Conclusions



Limites
Boîte noire
Peu explicable

Application du machine learning :  
**Prédiction du potentiel de croissance**, exemple  
sur les versements


Introduction

1. Valeur client

2. Clustering


**3. Prédiction des flux futurs**

Conclusions



### Limites


- Boîte noire
- Peu explicable



### Solution


Valeurs de Shapley

- Introduction
- 1. Valeur client
- 2. Clustering
- 3. Prédiction des flux futurs**
- Conclusions



### Limites

- Boîte noire
- Peu explicable



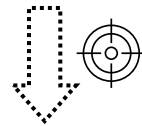
### Solution

- Valeurs de Shapley

## Pourquoi utiliser les valeurs de Shapley ?



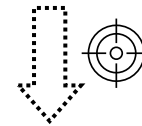
**Interpréter** les prédictions pour **chaque client**



**Décomposer** les scores de manière individuelle



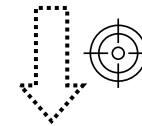
**Interpréter** les prédictions pour **l'ensemble des clients**



Dégager des **typologies** de **clients**



**Valider** la **cohérence** du modèle



Renforce la **transparence** et la **confiance** dans le modèle.

# Valeurs de Shapley | Interprétation pour 3 types de profils

Introduction

1. Valeur client

2. Clustering

**3. Prédiction des flux futurs**

Conclusions



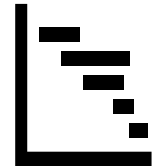
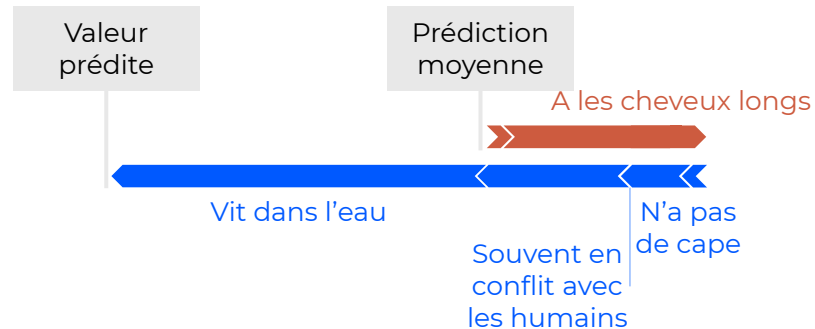
Superman



Wonder Woman



Aquaman



En combinant les décompositions individuelles des scores, il est possible d'en conclure des interprétations globales.

Introduction

1. Valeur client
2. Clustering
3. Prédiction des flux futurs

**Conclusions**

# Conclusions

---

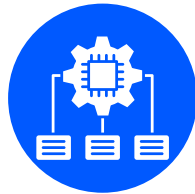
Introduction

1. Valeur client

2. Clustering

3. Prédiction des flux futurs

**Conclusions**



Utilisation du  
**machine learning**

**Différentes utilisations du machine learning** pour améliorer notre valeur client

**Des modèles supervisés** (prédiction) et **non supervisés** (clustering)

**Permet de mieux comprendre la valeur client**



Suivi des  
**performances des modèles**

**Indicateurs de suivi et de performance** de nos modèles

**Backtesting**



**Des modèles évolutifs** pour une amélioration continue

**Ajustement permanent**

**Enrichir nos modèles** avec de **nouvelles données** (open data) mais aussi par **réapprentissage** d'année en année.

**Prendre en compte l'impact des stratégies marketing ciblées dans nos modèles**

# Merci

---