

# Maîtriser la Prédiction : Classification vs Régression

LAYIBE YAYIBE Narcisse

27 juillet 2024

# Introduction

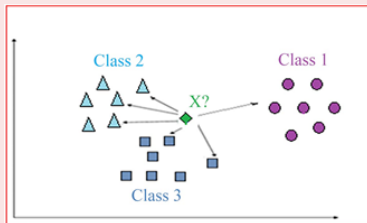
Les algorithmes de classification et de régression sont les piliers de l'apprentissage automatique, transformant les données en informations exploitables pour résoudre des problèmes complexes. Ils offrent des solutions précises et efficaces à une large gamme de défis, de la prédiction de catégories à l'estimation de valeurs continues.

# Définitions

## Classification & Régression

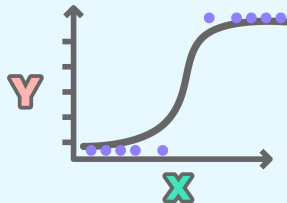
### Classification

- Prédiction de variables discrètes, telles que des catégories ou des étiquettes, à partir de données d'entrée.



### Régression

- Prédiction de variables continues, comme des valeurs numériques, à partir de données d'entrée.



# Comparaison des Algorithmes

## Classification et Régression

Critère	Classification	Régression
<b>Types d'algorithmes</b>	SVM, KNN, Naive Bayes	Régression linéaire, Réseaux de neurones
<b>Sorties</b>	Catégories/Étiquettes	Valeurs continues
<b>Applications</b>	Filtrage de spam, Diagnostic médical	Prévisions de prix, Analyse financière
<b>Avantages</b>	Interprétabilité, flexibilité	Précision, capacité de généralisation
<b>Inconvénients</b>	Sensibilité aux données, complexité	Besoin de données nombreuses, calcul intensif

# Problèmes résolus (1)

Par les algorithmes de classification

- 1 **Reconnaissance d'image** : Identifier la présence ou le type d'objet dans une image.
- 2 **Décisions binaires** : Prédire si un événement se produira ou non, comme l'acceptation ou le rejet d'un prêt.
- 3 **Segmentation de marché** : Classer les clients en groupes homogènes pour une meilleure stratégie marketing.
- 4 **Filtrage de spam** : Classer les emails en "spam" ou "non-spam" pour protéger les boîtes de réception.
- 5 **Détection de fraude** : Identifier les transactions frauduleuses dans les paiements par carte bancaire.
- 6 **Diagnostic médical** : Classer les symptômes pour aider les médecins à poser un diagnostic.
- 7 **Reconnaissance vocale** : Transcrire la parole en texte ou identifier les locuteurs.
- 8 **Analyse des sentiments** : Déterminer le ton ou l'attitude dans les textes, par exemple, positif, négatif ou neutre.
- 9 **Recommandation de produits** : Classer les utilisateurs dans des groupes pour suggérer des produits pertinents.
- 10 **Détection d'anomalies** : Identifier les comportements inhabituels ou suspects dans les réseaux informatiques ou les systèmes financiers.

# Problèmes résolus (2)

Par les algorithmes de régression

- ➊ **Prévisions de ventes** : Estimer les ventes futures en fonction de facteurs tels que le prix, la publicité et la concurrence.
- ➋ **Évaluation de risque** : Prédire le risque de défaut de paiement ou de faillite d'une entreprise.
- ➌ **Diagnostic médical** : Prédire la progression d'une maladie ou l'évolution d'un traitement à partir de données médicales.
- ➍ **Prévisions de prix** : Estimer le prix de l'immobilier en fonction de la superficie, du nombre de pièces, de la localisation, etc.
- ➎ **Recommandations personnalisées** : Prédire les préférences et les intérêts des utilisateurs pour faire des recommandations pertinentes.
- ➏ **Optimisation de la production** : Prédire la quantité optimale de matières premières à commander pour minimiser les coûts.
- ➐ **Analyse financière** : Prédire la performance future des actions ou des portefeuilles financiers.
- ➑ **Prévisions météorologiques** : Prédire les conditions météorologiques futures en fonction des données historiques et des modèles climatiques.
- ➒ **Gestion de l'énergie** : Estimer la consommation future d'énergie pour optimiser la production et la distribution.
- ➓ **Prévision de la demande** : Estimer la demande future de produits ou services pour une meilleure gestion des stocks et de la logistique.

# Exemples d'applications (1)

Pour les algorithmes de classification

- ❶ **Filtrage de spam** : Classer les emails en "spam" ou "non-spam" pour protéger les boîtes de réception.
- ❷ **Détection de fraude** : Identifier les transactions frauduleuses dans les paiements par carte bancaire.
- ❸ **Diagnostic médical** : Classer les symptômes pour aider les médecins à poser un diagnostic.
- ❹ **Reconnaissance d'image** : Identifier les objets, personnes ou animaux dans des images.
- ❺ **Analyse des sentiments** : Déterminer le ton ou l'attitude dans les textes, par exemple, positif, négatif ou neutre.
- ❻ **Systèmes de recommandation** : Classer les produits ou services pour suggérer des recommandations pertinentes.
- ❼ **Sécurité informatique** : Détecter les intrusions ou les activités suspectes sur un réseau.
- ❽ **Reconnaissance vocale** : Transcrire la parole en texte ou identifier les locuteurs.
- ❾ **Analyse de documents** : Classer des documents en différentes catégories, par exemple, par sujet ou type.
- ❿ **Classification des clients** : Segmenter les clients en groupes homogènes pour des stratégies marketing ciblées.

# Exemples d'applications (2)

Pour les algorithmes de régression

- 1 **Prévisions de prix** : Estimer le prix de l'immobilier en fonction de la superficie, du nombre de pièces, de la localisation, etc.
- 2 **Recommandations personnalisées** : Prédire les préférences et les intérêts des utilisateurs pour faire des recommandations pertinentes.
- 3 **Optimisation de la production** : Prédire la quantité optimale de matières premières à commander pour minimiser les coûts.
- 4 **Prévisions de ventes** : Estimer les ventes futures en fonction de facteurs tels que le prix, la publicité et la concurrence.
- 5 **Évaluation de risque** : Prédire le risque de défaut de paiement ou de faillite d'une entreprise.
- 6 **Diagnostic médical** : Prédire la progression d'une maladie ou l'évolution d'un traitement à partir de données médicales.
- 7 **Analyse financière** : Prédire la performance future des actions ou des portefeuilles financiers.
- 8 **Prévisions météorologiques** : Prédire les conditions météorologiques futures en fonction des données historiques et des modèles climatiques.
- 9 **Gestion de l'énergie** : Estimer la consommation future d'énergie pour optimiser la production et la distribution.
- 10 **Prévision de la demande** : Estimer la demande future de produits ou services pour une meilleure gestion des stocks et de la logistique.



# Avantages et inconvénients de chaque type d'algorithme

## Avantages de la classification

- Résolution de problèmes binaires ou à classes multiples.
- Interprétabilité des résultats.
- Adaptation à une grande variété de domaines.

## Avantages de la régression

- Prédiction de valeurs continues.
- Prise en compte de la complexité des relations entre variables.
- Possibilité de généralisation à de nouvelles données.

## Inconvénients

- Sensibilité aux données d'entraînement.
- Difficulté à traiter les variables non linéaires.
- Besoin d'une bonne compréhension des hypothèses sous-jacentes.

# Conclusion & Recommandations

- 1 **Choisir l'algorithme adapté** : Identifier le type de problème à résoudre (classification ou régression) est la première étape cruciale.
- 2 **Comprendre les données** : Analyser attentivement les caractéristiques des données pour sélectionner les bons algorithmes et paramètres.
- 3 **Évaluer les performances** : Utiliser des métriques pertinentes pour mesurer la qualité des prédictions et itérer sur les modèles.

Merci de votre  
attention!



# Restons en contact

Connectons-nous pour continuer la conversation

- **LinkedIn** : <https://www.linkedin.com/in/layibe-yayibe-narcisse-ba2a34281/>
- **Github** : [github.com/layibe-02/](https://github.com/layibe-02/)
- **Email** : [narcisse.layibe03@gmail.com](mailto:narcisse.layibe03@gmail.com)
- **Email** : [narcisse.layibe@facsciences-uy1.cm](mailto:narcisse.layibe@facsciences-uy1.cm)