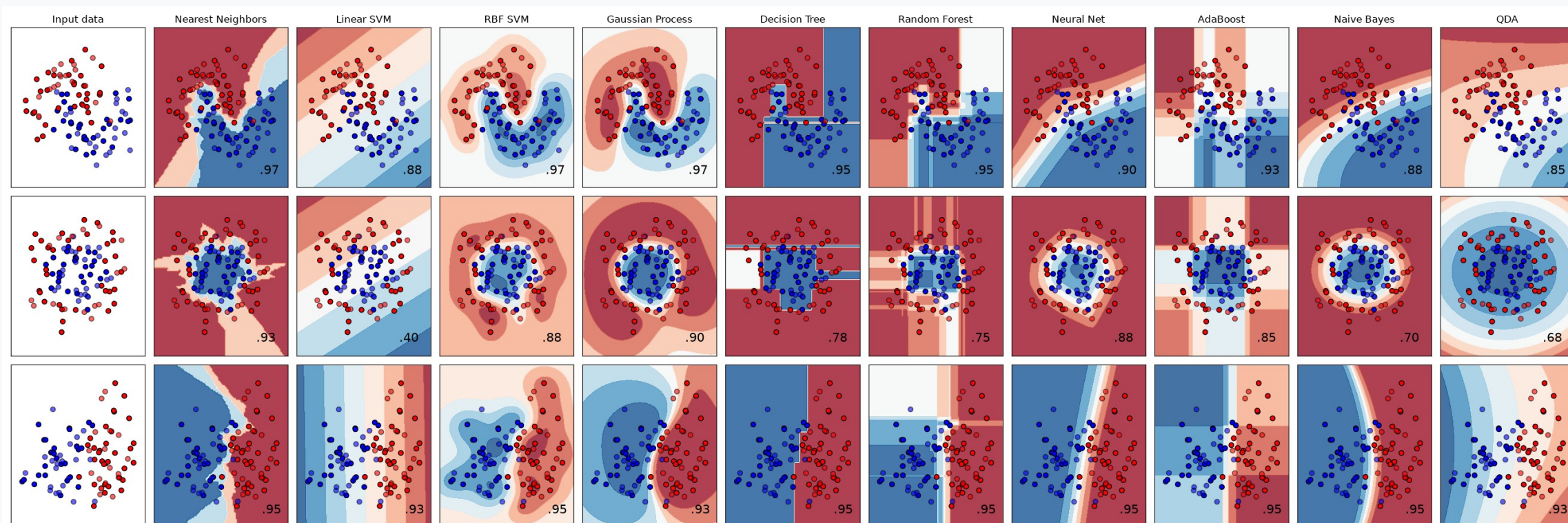


Sept familles de modèles

Même problème, hypothèses et compromis différents.

Trois jeux 2D, des frontières différentes



Voisins · SVM · arbre · forêt · réseau · boosting : comparer les formes, pas les scores.

Illustration scikit-learn, jeux synthétiques 2D · BSD-3-Clause

Modèles linéaires · combiner les variables

IDÉE

Une somme pondérée des variables produit un score.

APPRENTISSAGE / PRÉDICTION

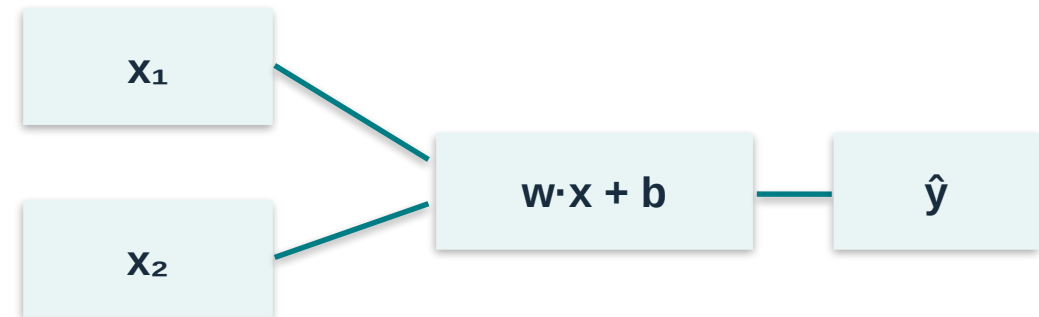
Ajuster les coefficients ; appliquer la même règle à chaque ligne.

À RÉGLER

Pénalisation : alpha (Ridge), C (logistique).

POINT DE VIGILANCE

Mettre à l'échelle ; relation non linéaire parfois mal captée.



Ridge : valeur numérique

Logistique : classe / probabilité

k-NN · apprendre par proximité

IDÉE

Une nouvelle observation hérite du vote de ses voisins.

APPRENTISSAGE / PRÉDICTION

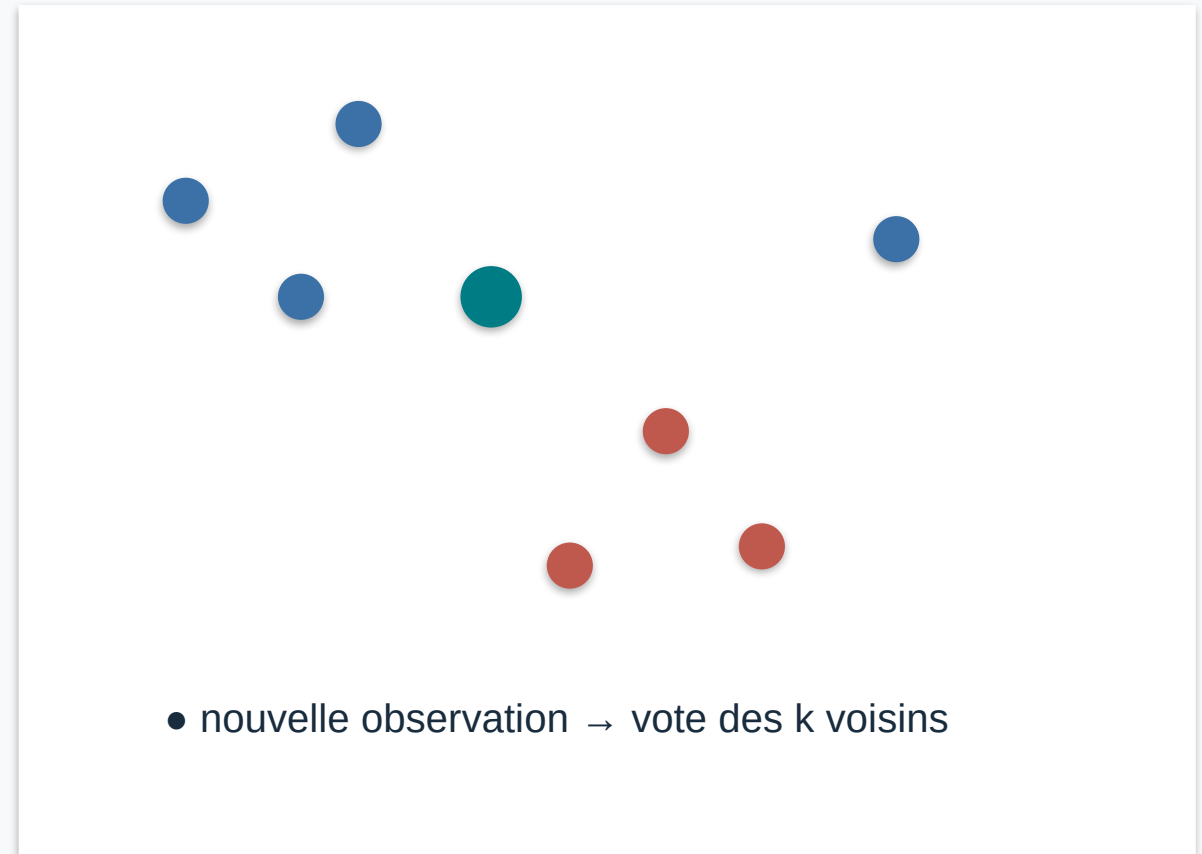
Conserver les exemples ; mesurer les distances au moment de prédire.

À RÉGLER

Nombre k, distance, pondération des voisins.

POINT DE VIGILANCE

Sensible à l'échelle ; prédiction coûteuse sur un grand jeu.



Arbre · enchaîner des questions

IDÉE

Des seuils découpent progressivement l'espace des variables.

APPRENTISSAGE / PRÉDICTION

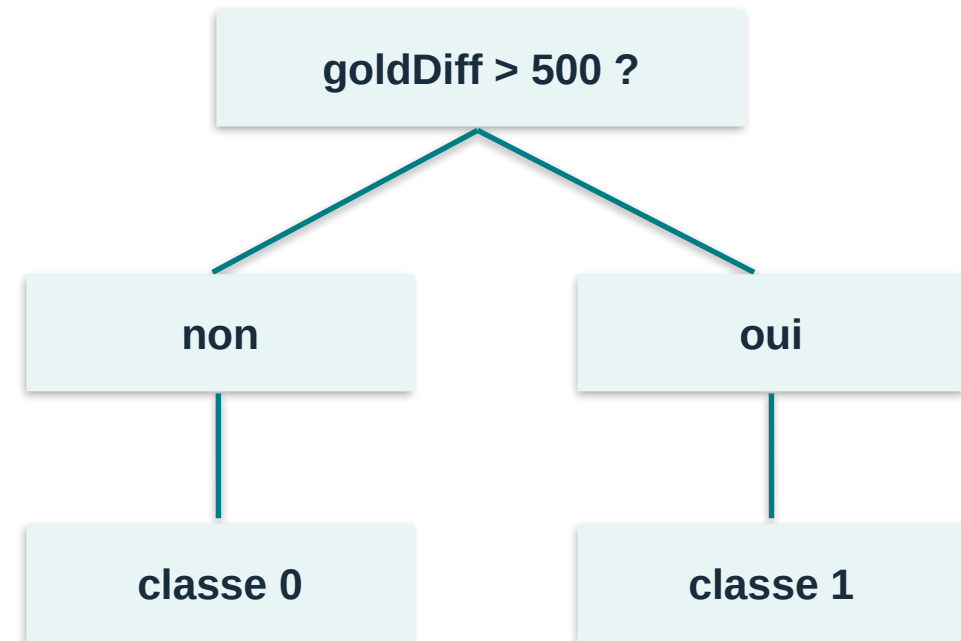
Choisir les coupures sur l'entraînement ; suivre une branche pour prédire.

À RÉGLER

Profondeur, taille minimale des feuilles.

POINT DE VIGILANCE

Arbre profond : mémorisation ; un petit changement peut modifier l'arbre.



Forêt aléatoire · agréger des arbres

IDÉE

Des arbres variés votent ou moyennent leurs réponses.

APPRENTISSAGE / PRÉDICTION

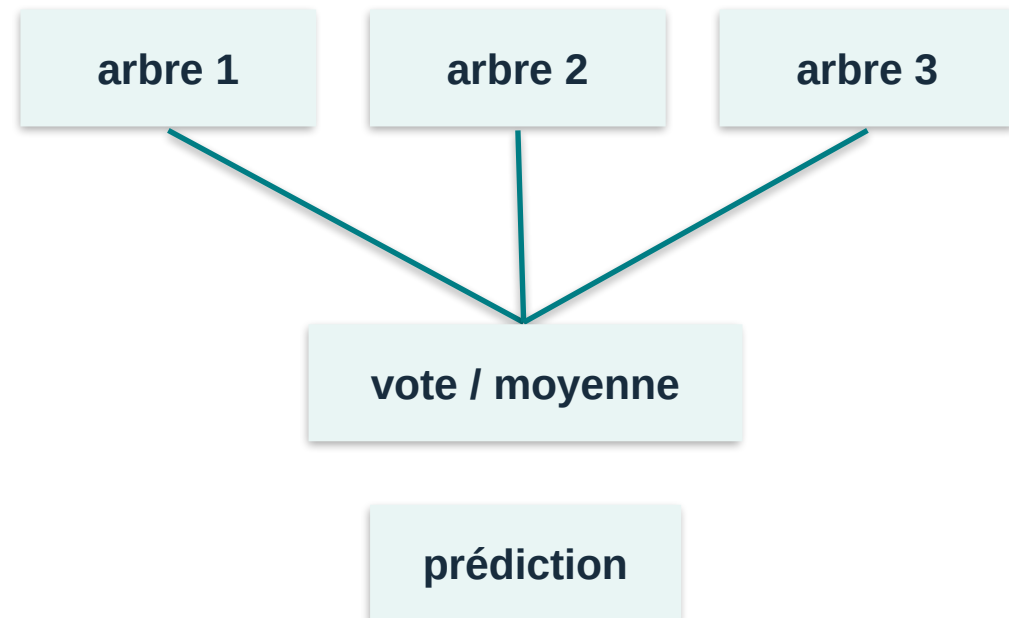
Chaque arbre voit un échantillon et des variables partiellement différents.

À RÉGLER

Nombre d'arbres, profondeur, taille des feuilles.

POINT DE VIGILANCE

Moins lisible qu'un arbre ; coût de calcul et mémoire supérieurs.



Boosting · construire séquentiellement

IDÉE

Chaque nouveau modèle corrige la perte de l'ensemble courant.

APPRENTISSAGE / PRÉDICTION

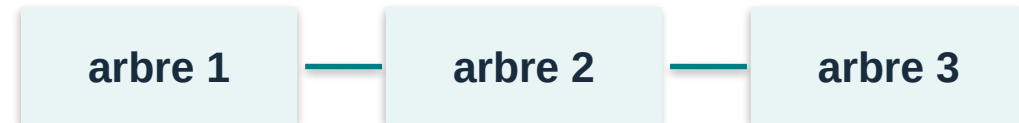
Ajouter des arbres successifs ; combiner leurs contributions.

À RÉGLER

Nombre d'itérations, profondeur, taux d'apprentissage.

POINT DE VIGILANCE

Peut surajuster ; surveiller validation et temps d'entraînement.



Chaque étape améliore une prédiction collective selon une fonction de perte.

SVM · séparer avec une marge

IDÉE

Chercher une frontière séparant les classes avec une grande marge.

APPRENTISSAGE / PRÉDICTION

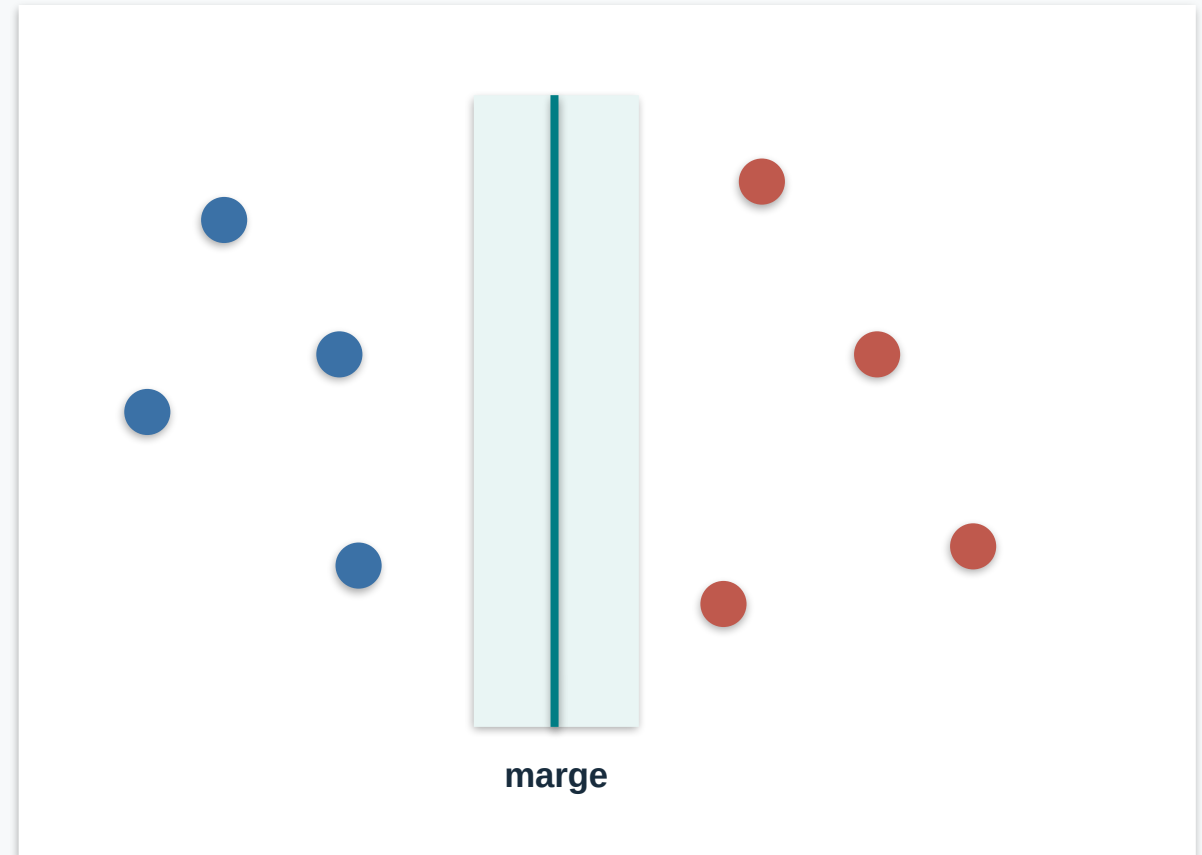
Optimiser la frontière ; un noyau peut créer une séparation non linéaire.

À RÉGLER

C et, selon le noyau, gamma.

POINT DE VIGILANCE

Échelle importante ; coût élevé sur très grands jeux de données.



Réseaux de neurones · couches apprises

IDÉE

Composer des transformations non linéaires entre entrée et sortie.

APPRENTISSAGE / PRÉDICTION

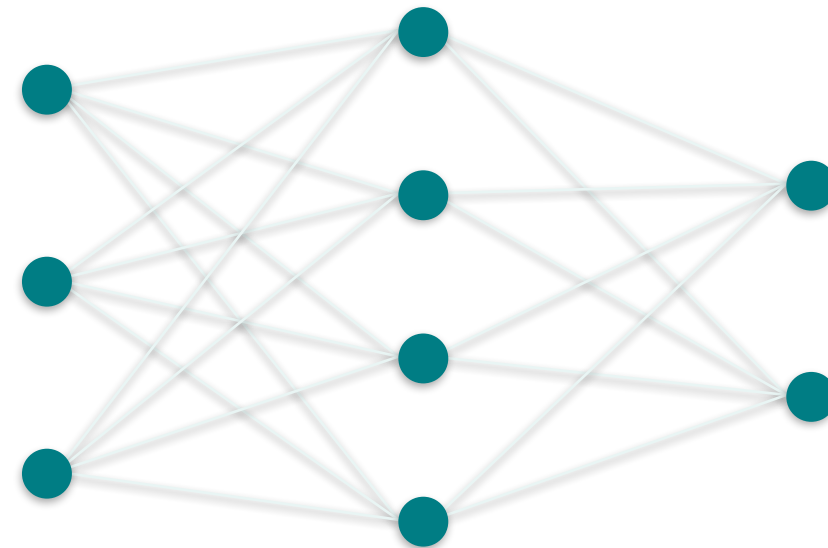
Ajuster les poids par optimisation itérative sur les données.

À RÉGLER

Architecture, régularisation, nombre d'itérations.

POINT DE VIGILANCE

Données, réglage et calcul ; MLP scikit-learn $\neq$ deep learning moderne.



entrées → couche cachée → sorties

Références et illustration

Familles de modèles — guide officiel

https://scikit-learn.org/stable/supervised_learning.html

Comparaison visuelle des classifieurs — scikit-learn

https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html

Droits sur les images : BSD-3-Clause — FAQ scikit-learn

<https://scikit-learn.org/stable/faq.html#how-can-i-obtain-permission-to-use-the-images-in-scikit-learn-for-my-work>