

Listes Algorithmes Importants en Machine Learning

Régression linéaire

- Utilisation : Modélisation de relations linéaires entre une variable dépendante et une ou plusieurs variables indépendantes.
- Fonctionnement : Minimisation de l'erreur quadratique moyenne pour ajuster une ligne droite aux données.

Régression logistique

- Utilisation : Classification binaire ou multiclasse.
- Fonctionnement : Modélisation de la probabilité d'appartenance à une classe en utilisant la fonction logistique.

Arbres de décision

- Utilisation : Classification et régression.
- Fonctionnement : Construction d'un arbre basé sur des règles de décision pour partitionner les données.

Forêts aléatoires

- Utilisation : Classification, régression, détection d'anomalies.
- Fonctionnement : Agrégation de plusieurs arbres de décision pour réduire le surajustement.

Machines à vecteurs de support (SVM)

- Utilisation : Classification et régression.
- Fonctionnement : Trouver l'hyperplan qui maximise la marge entre les classes.

K-moyennes

- Utilisation : Clustering.
- Fonctionnement : Partitionnement des données en K clusters en minimisant la distance entre les points et les centres de cluster.

K-plus proches voisins (K-NN)

- Utilisation : Classification et régression.
- Fonctionnement : Prédiction basée sur la majorité des K voisins les plus proches dans l'espace des caractéristiques.

Analyse en composantes principales (PCA)

- Utilisation : Réduction de dimension.
- Fonctionnement : Transformation linéaire pour réduire la dimension tout en conservant la variance maximale.

Naïve Bayes

- Utilisation : Classification.
- Fonctionnement : Utilisation du théorème de Bayes pour estimer les probabilités conditionnelles.

Régression polynomiale

- Utilisation : Modélisation de relations non linéaires.
- Fonctionnement : Utilisation de polynômes pour ajuster les données.

Réseau bayésien

- Utilisation : Modélisation de dépendances probabilistes entre variables.
- Fonctionnement : Représentation graphique des relations entre les variables.

K-means++

- Utilisation : Clustering.
- Fonctionnement : Une amélioration de l'algorithme K-moyennes qui initialise les centroïdes de manière plus efficace.

Isolation Forest

- Utilisation : Détection d'anomalies.
- Fonctionnement : Crée un arbre de décision pour isoler des données anormales plus rapidement que les données normales.

Algorithmes génétiques

- Utilisation : Optimisation, sélection de caractéristiques.
- Fonctionnement : Utilise des opérateurs génétiques (croisement, mutation) pour trouver des solutions optimales.

Perceptron Multicouche (MLP)

- **Utilisation** : Classification, régression, approximation de fonctions.
- **Fonctionnement** : Réseau de neurones avec une ou plusieurs couches cachées.

Machines à vecteurs de support à noyau (SVM à noyau)

- **Utilisation** : Classification et régression non linéaires.
- **Fonctionnement** : Extension des SVM pour gérer des données non linéaires en utilisant des noyaux.

Méthodes ensemblistes (AdaBoost, Gradient Boosting, XGBoost)

- **Utilisation** : Amélioration de la performance des modèles de base.
- **Fonctionnement** : Combinaison de modèles faibles pour créer un modèle plus puissant.

Méthodes de réduction de dimension (t-SNE, LLE, UMAP)

- Utilisation : Visualisation de données à haute dimension.
- Fonctionnement : Projection des données dans un espace de dimension réduite.

Régression de Poisson

- Utilisation : Modélisation de données de comptage.
- Fonctionnement : Modèle de régression adapté pour les données de comptage.