



5

Modèles Ensemblistes

5. Modèles – Ensemblistes



apprentissage automatique dans lequel plusieurs modèles (souvent appelés "apprenants faibles") sont formés pour résoudre le même problème et combinés pour obtenir de meilleurs résultats (plus précis et/ou robuste).

Attention : combinaison des modèles apprenants faibles à faibles biais et fortes variances → réduire la variance (et vis et versa)

Méthodes parallèles : 

entraîner plusieurs modèles apprenants faibles de manière indépendante (en parallèle) pour ensuite les regrouper afin de prendre une décision

Algorithme de combinaison ?

BAGGING:

Entraîne plusieurs modèles apprenants faibles homogènes, indépendamment les uns des autres en **parallèle** et les combine en suivant un processus déterministe de calcul de moyenne (moins de variance).

BOOSTING:

Entraîne des modèles apprenants faibles homogènes **séquentiellement** de manière très adaptative (un modèle de base dépend des précédents) et les combine en suivant une stratégie déterministe (modèles forts moins biaisés).

Méthodes séquentielles : 

entraîner plusieurs modèles apprenants faibles de manière séquentielle à la suite, premier modèle induit erreurs, on donne plus de poids aux données induisant erreur et entraîne deuxième modèle, puis nouveau modèle entraîné avec le jeu de données pondérés en augmentant le poids des observations induisant le plus d'erreur... Le dernier modèle est le plus performant.

STACKING:

Entraîne des modèles apprenants faibles hétérogènes en **parallèle** et les combine en formant un méta-modèle pour produire une prédiction basée sur les prédictions des différents modèles faibles (modèles forts moins biaisés).