**3. Partie II – Réseaux de neurones de base**

La deuxième partie se consacre aux architectures les plus simples et aux mécanismes d’apprentissage fondamentaux.

* **Perceptron multicouche (MLP)** : on y explique comment empiler des couches de neurones artificiels pour capturer des relations non linéaires. Le rôle des fonctions d’activation – qui introduisent la non-linéarité – est détaillé, tout comme les stratégies d’initialisation des poids pour assurer une convergence rapide.
* **Rétropropagation** : le mécanisme central d’apprentissage est présenté pas à pas. Le texte décrit comment le gradient du critère d’erreur est calculé pour chaque paramètre en remontant du signal de sortie vers les couches d’entrée.
* **Optimisation stochastique** : différentes variantes de descente de gradient sont comparées – de la forme la plus simple à des algorithmes adaptatifs qui ajustent automatiquement la vitesse d’apprentissage en fonction de l’historique des gradients. Les techniques de réglage du taux d’apprentissage, comme les décroissances graduelles ou les cycles périodiques, sont expliquées.
* **Régularisation** : pour éviter le surapprentissage, plusieurs techniques sont décrites : la pénalisation de la taille des poids, la désactivation aléatoire de neurones pendant l’entraînement, ainsi que la normalisation des activations. La section couvre aussi la stratégie d’arrêt anticipé et l’importance de la validation croisée.