

# Validation

On veut avoir un modèle entraîné et une idée de ses performances.

## 🔗 Validation set method

On utilise une partie du training pour la validation.

C'est simple et assez efficace.  
mais on perd des données.

## 🔗 Leave-One-Out (LOO) Cross Validation

On fait **N entraînements**, où N est le nombre d'exemples dans le dataset :

- À chaque fois, on **laisse un exemple pour la validation**, et on entraîne sur les  $N - 1$  autres.
- Puis on moyenne les erreurs de validation.

On utilise **quasiment toutes les données** pour l'entraînement à chaque étape, c'est très **précis** pour estimer la performance  
mais **coûteux en temps de calcul**, surtout pour grands jeux de données

## 🔗 Le mieux : k-fold cross validation

On sépare le dataset en k parties, on entraîne avec k-1 et on teste avec la dernière.