

# Mathématiques pour l'intelligence artificielle

M2 informatique parcours IAAA, Aix-Marseille Université

Thomas Schatz ([thomas.schatz@univ-amu.fr](mailto:thomas.schatz@univ-amu.fr))

Mardi 2 septembre 2025

# Questions pratiques

- Calendrier des cours sur le site IAAA + ENT AMU
- Questions, etc. mattermost M2 IAAA, canal MIA (coming soon?)
- Matériel de cours sur mon site web (<https://thomas.schatz.cogserver.net/teaching/>)
- Evaluation: Note finale =  $\text{Max}(0.5 \cdot \text{CC} + 0.5 \cdot \text{ET}; \text{ET})$ 
  - CC contrôle continu:
    - 1 point : scribe pour une séance
    - 2 points : exercices à rendre (correction d'une sélection aléatoire)
    - 7 points : partiel, mardi 23 septembre 10h15-12h15
  - ET examen terminal: mardi 14 Octobre 10h15-12h15
- Partiel et examen terminal:
  - Ecrit, document manuscrits uniquement, pas d'ordinateur, tablette, calculette, montre, lunettes ou tout autre objet connecté...

# Scribes

- Document à me rendre au plus tard une semaine après la séance (rendre un seul document par groupe)
- 2/09 Clara Karkach, Jules Danielou
- 9/09 Mohamed Alloui, Cherifi Aymen
- 16/09 Raissa Mezine, Sajjad Ghasemian
- 24/09 Joe Hubi, Mohamed Alloui
- 25/09 Idir Bellili, Saal Racim
- 30/09 Anlaoudine Moindze Boina
- 7/10 Amal Bourahma Atchory Rameaux Mel

# Plan du cours (prévisionnel)

Introduction générale

1. Algèbre linéaire (2h)

2. Optimisation (4h)

Partiel

3. Modélisation probabiliste (2h)

4. Probabilités, statistique et théorie de l'apprentissage (4h)

Terminal

# Introduction générale

# Introduction générale

**Comment définir l'apprentissage automatique ?**

# Introduction générale

Comment définir l'apprentissage automatique ?

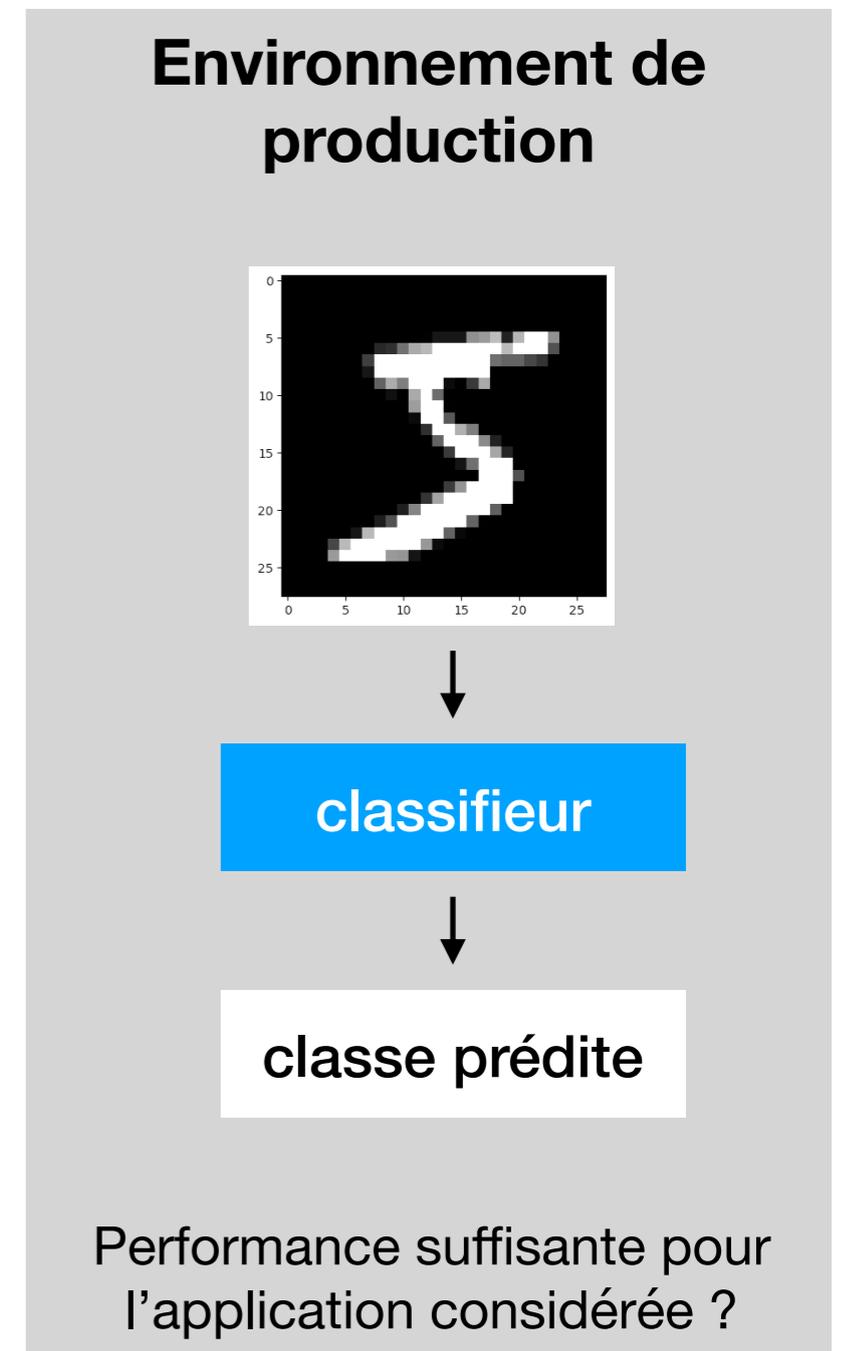
**Généralisation**

**+**

**Calcul intensif**

# Introduction générale

Structure typique d'un problème d'apprentissage automatique

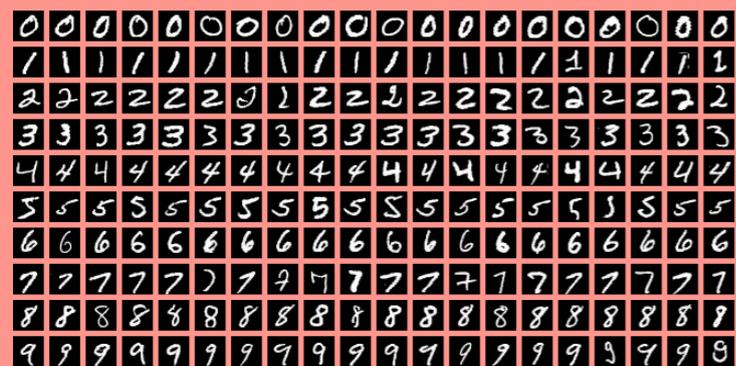


# Introduction générale

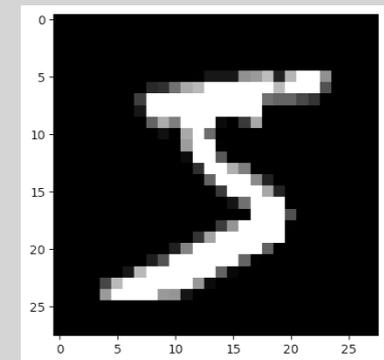
## Structure typique d'un problème d'apprentissage automatique

### Entraînement (procédure de recherche)

#### Données d'entraînement



### Environnement de production



classifieur

classe prédite

Performance suffisante pour l'application considérée ?

# Introduction générale

## Structure typique d'un problème d'apprentissage automatique

### Entraînement (procédure de recherche)

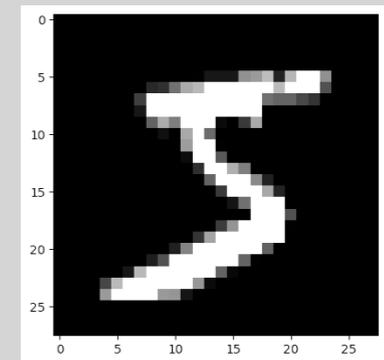
Espace de recherche



### Données d'entraînement

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

### Environnement de production



classifieur

classe prédite

Performance suffisante pour l'application considérée ?

# Introduction générale

## Structure typique d'un problème d'apprentissage automatique

### Entraînement (procédure de recherche)

Espace de recherche

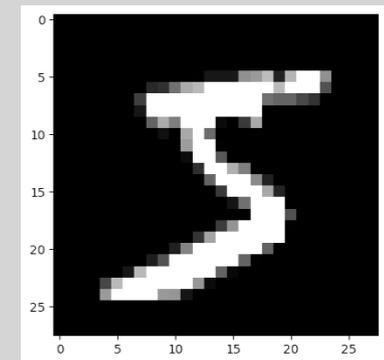


### Données d'entraînement

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

Fonction de perte

### Environnement de production



classifieur

classe prédite

Performance suffisante pour l'application considérée ?

# Introduction générale

## Structure typique d'un problème d'apprentissage automatique

### Entraînement (procédure de recherche)

Espace de recherche

Hypothèse/paramètre

Données d'entraînement

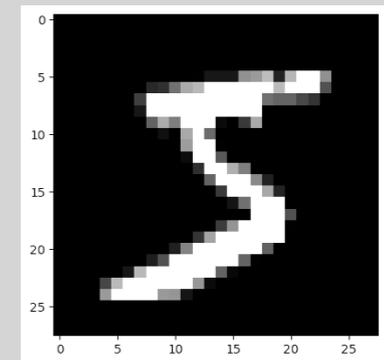
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

Perte

Fonction de perte

Evaluation d'une hypothèse

### Environnement de production



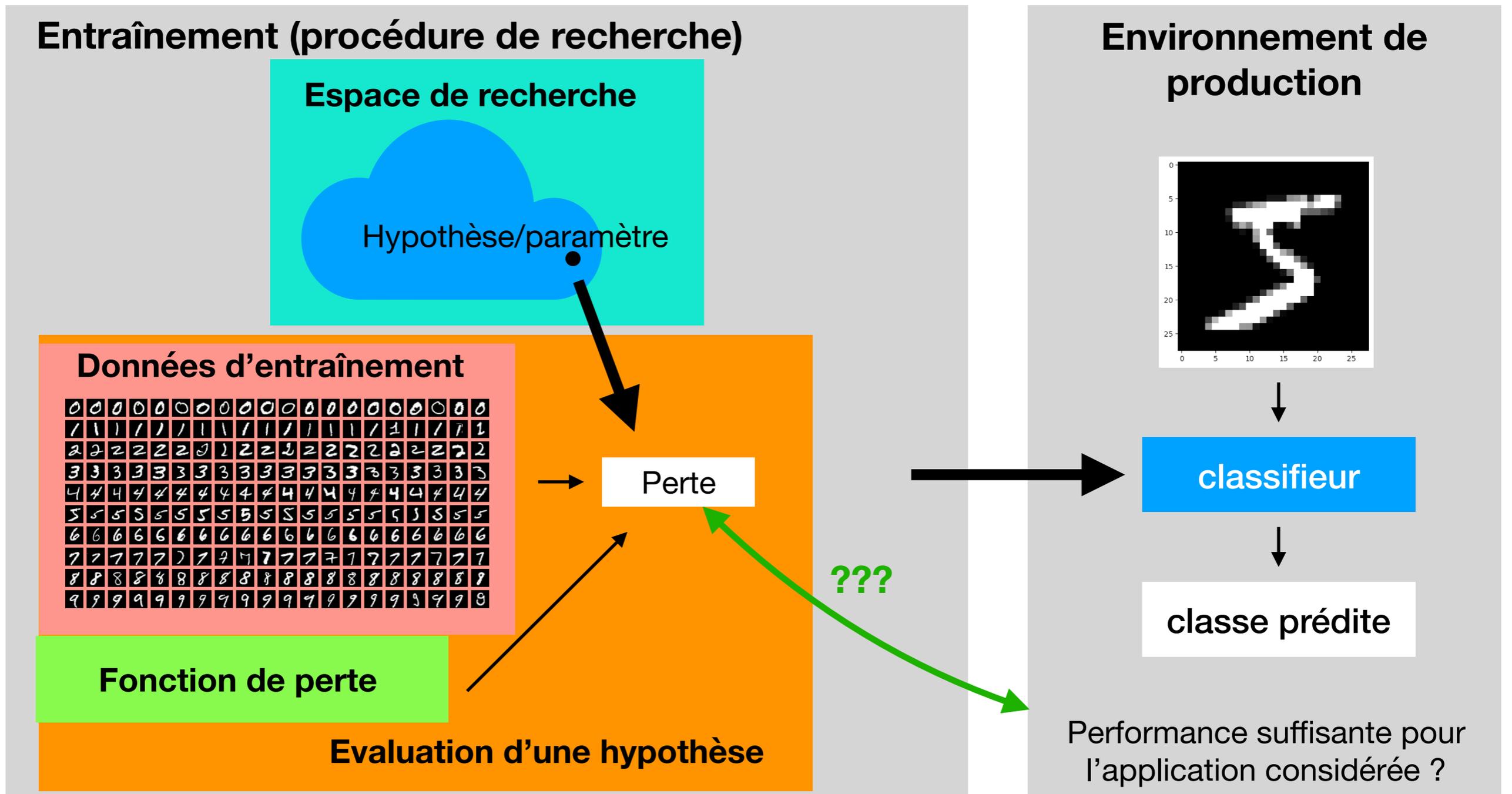
classifieur

classe prédite

Performance suffisante pour l'application considérée ?

# Introduction générale

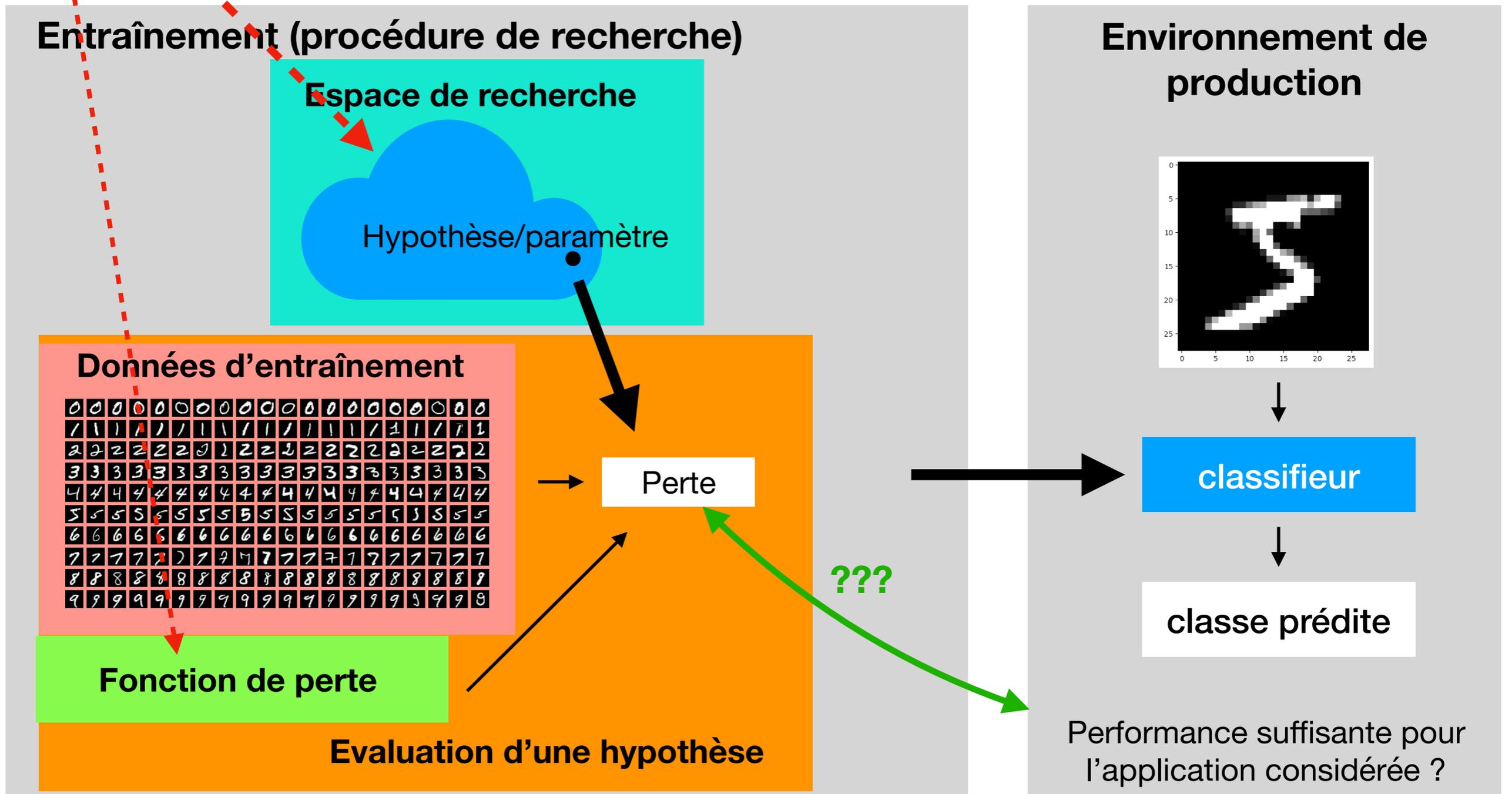
## Structure typique d'un problème d'apprentissage automatique



# Introduction générale

Algèbre linéaire

Structure typique d'un problème d'apprentissage automatique



# Introduction générale

Optimisation

Structure typique d'un problème d'apprentissage automatique

Entraînement (**procédure de recherche**)

Espace de recherche

Hypothèse/paramètre

Données d'entraînement

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

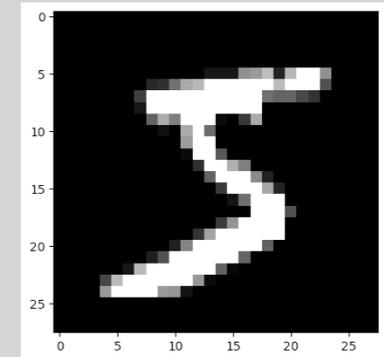
Fonction de perte

Perte

???

Evaluation d'une hypothèse

Environnement de production



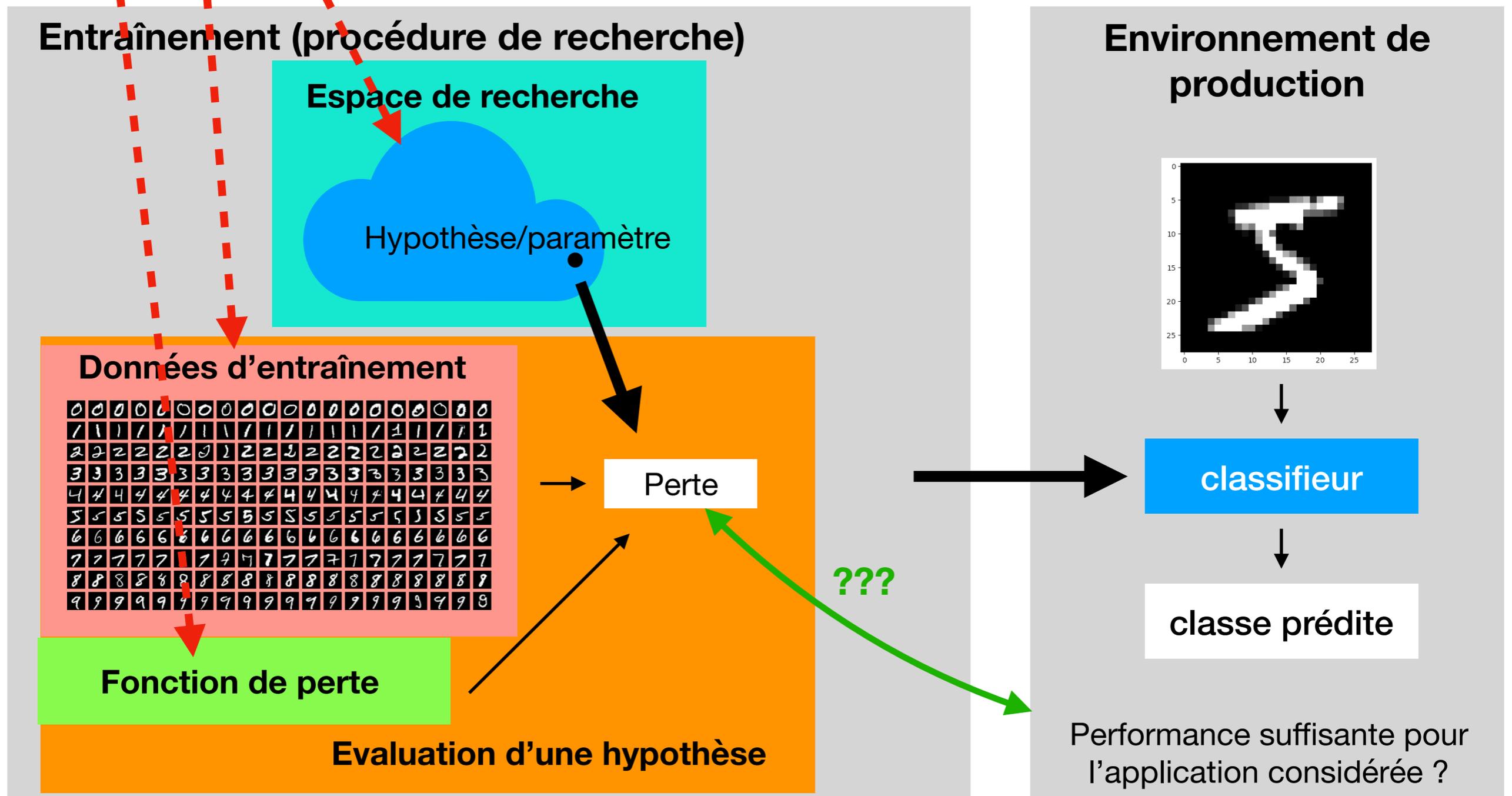
classifieur

classe prédite

Performance suffisante pour l'application considérée ?

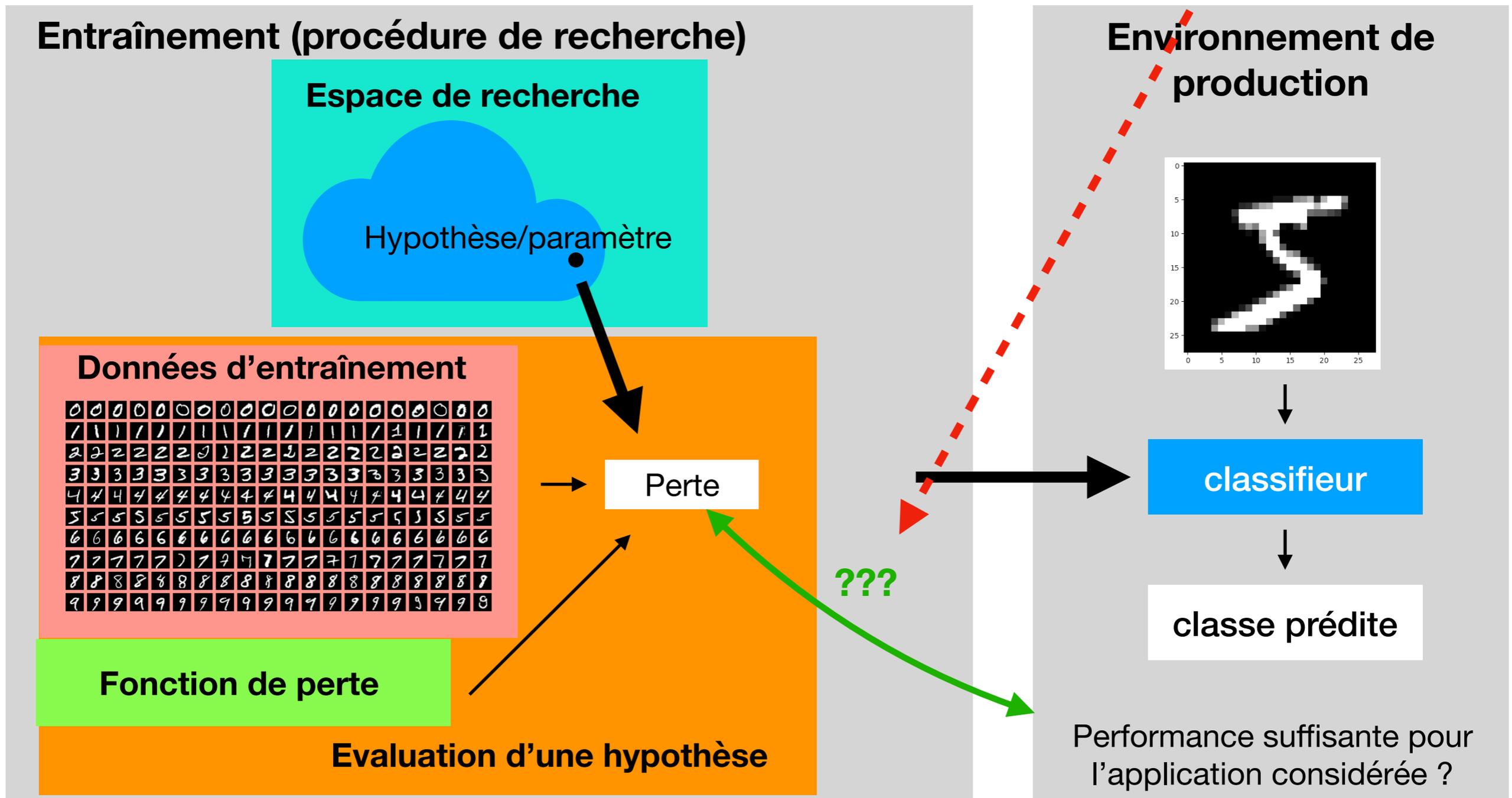
# Introduction générale

Structure typique d'un problème d'apprentissage automatique



# Introduction générale

Structure typique d'un problème d'apprentissage automatique



# Algèbre linéaire

# Algèbre linéaire

- Plusieurs interprétations des opérations matricielles
- Produit scalaire, orthogonalité, normes
- Décomposition en valeurs singulières
- Théorème spectral, positivité, déterminant, trace

# Algèbre linéaire

- Plusieurs interprétations des opérations matricielles
- Produit scalaire, orthogonalité, normes
- Décomposition en valeurs singulières
- Théorème spectral, positivité, déterminant, trace

# Interpretation des opérations matricielles

Multiplication matrice/vecteur colonne:  $(Mv)_i := \sum_{j=1}^n m_{ij} v_j$

Multiplication matrice-matrice:  $(AB)_{ij} := \sum_{k=1}^n a_{ik} b_{kj}$

# Interpretation des opérations matricielles

## Notation par bloc

La matrice

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 & 2 \\ 3 & 3 & 4 & 4 & 4 \\ 3 & 3 & 4 & 4 & 4 \\ 3 & 3 & 4 & 4 & 4 \end{bmatrix}$$

peut être partitionnée en quatre blocs

$$\mathbf{P}_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{P}_{12} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}, \mathbf{P}_{21} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \\ 3 & 3 \end{bmatrix}, \mathbf{P}_{22} = \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}.$$

On peut alors écrire la matrice par bloc comme :

$$\mathbf{P}_{\text{partitionnee}} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}.$$

# Interpretation des opérations matricielles

write  $A \in \mathbf{R}^{m \times n}$  in terms of its columns:

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

then  $y = Ax$  can be written as

$$y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

write  $A$  in terms of its rows:

$$A = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_n^T \end{bmatrix}$$

then  $y = Ax$  can be written as

$$y = \begin{bmatrix} \tilde{a}_1^T x \\ \tilde{a}_2^T x \\ \vdots \\ \tilde{a}_m^T x \end{bmatrix}$$

# Interpretation des opérations matricielles

$$c_{ij} = \tilde{a}_i^T b_j = \langle \tilde{a}_i, b_j \rangle$$

$$C = [ c_1 \cdots c_p ] = AB = [ Ab_1 \cdots Ab_p ]$$

$$C = \begin{bmatrix} \tilde{c}_1^T \\ \vdots \\ \tilde{c}_m^T \end{bmatrix} = AB = \begin{bmatrix} \tilde{a}_1^T B \\ \vdots \\ \tilde{a}_m^T B \end{bmatrix}$$

$$C = \sum_i a_i \tilde{b}_i^T$$

# Algèbre linéaire

- Plusieurs interprétations des opérations matricielles
- Produit scalaire, orthogonalité, normes
- Décomposition en valeurs singulières
- Théorème spectral, positivité, déterminant, trace

# Produit scalaire usuel et norme euclidienne

$$(a | b) = \sum_{i=1}^n a_i b_i = a^T b = \|a\|_2 \|b\|_2 \cos(\theta)$$

$$\|x\|_2 = \sqrt{x^T x} = \sqrt{(x|x)}$$

# Angles et orthogonalité

## Intérêt ?

Soit  $E$  un espace vectoriel de dimension finie  $n$  muni d'un produit scalaire  $(\cdot | \cdot)$  et  $V = (v_1, \dots, v_n)$  une base orthonormale de  $E$ . Alors, pour tout  $v \in E$ ,

$$v = \sum_{i=1}^n (v | v_i) v_i.$$

# Algèbre linéaire

- Plusieurs interprétations des opérations matricielles
- Produit scalaire, orthogonalité, normes
- Décomposition en valeurs singulières
- Théorème spectral, positivité, déterminant, trace

Supposez que vous avez un réseau de neurone avec la matrice de poids

$$W = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

pour la couche L. Comment comprendre comment les activations de la couche L-1 sont transformées au niveau de la couche L dans ce réseau ?

# Décomposition en valeurs singulières

$$AV = U\Sigma \quad A \begin{bmatrix} v_1 & \dots & v_r & \dots & v_n \end{bmatrix} = \begin{bmatrix} u_1 & \dots & u_r & \dots & u_m \end{bmatrix} \left[ \begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ \hline & & & 0 \end{array} \right]$$