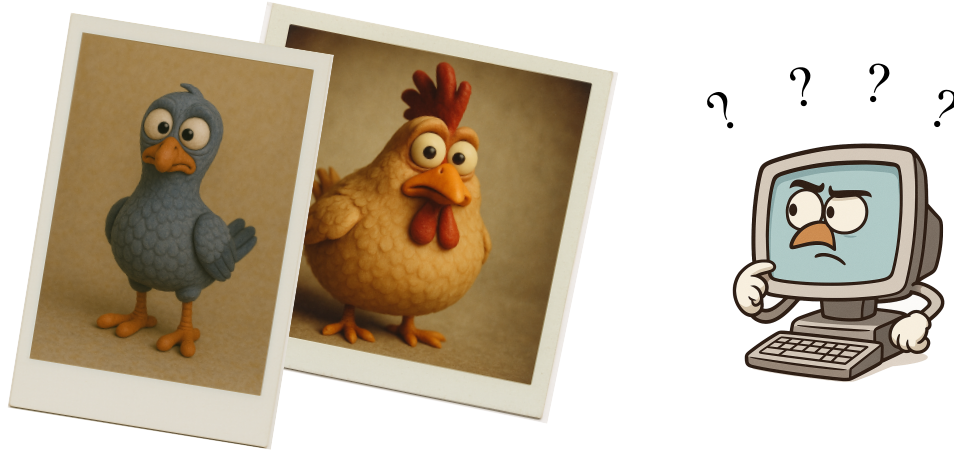


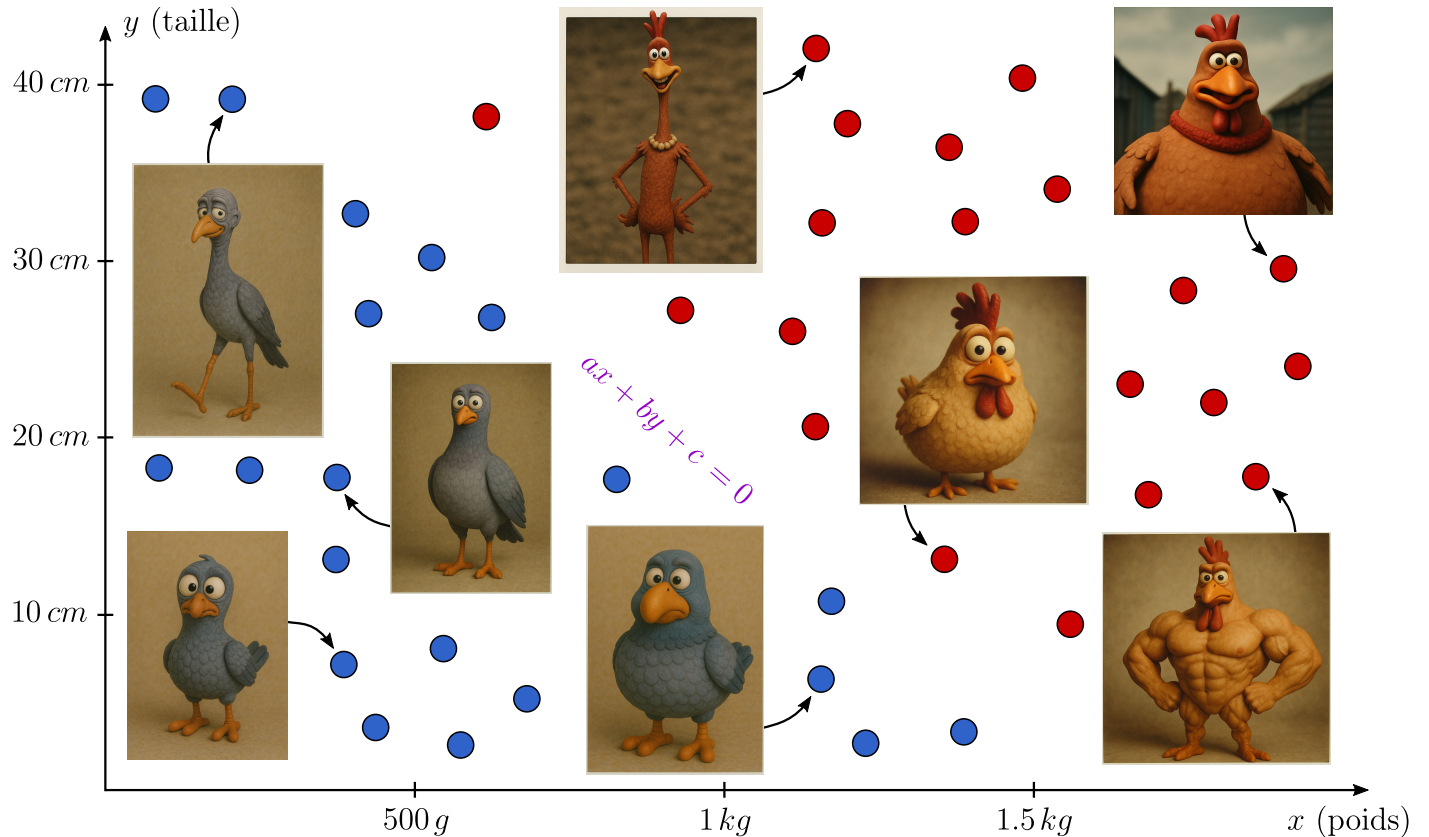
Comment une machine apprend à reconnaître les oiseaux ? (et bien d'autres choses...)

1 La problématique

Comment un ordinateur peut-il apprendre à distinguer deux types d'oiseaux — par exemple des poulets et des pigeons — à partir de leurs caractéristiques physiques (comme la taille et le poids) ?



Chaque oiseau est représenté par **un point dans le plan** dont les coordonnées traduisent ses mesures (poids, taille). Ainsi, l'ensemble des individus forme **deux nuages de points** : celui des poulets et celui des pigeons.



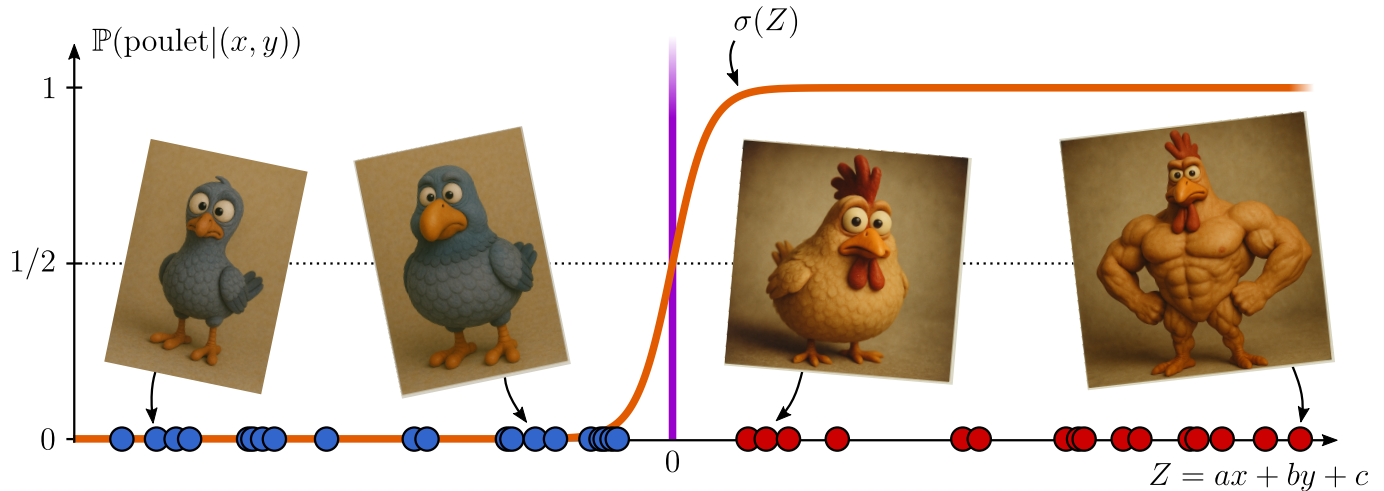
2 Idée de la régression logistique

On cherche à modéliser la **probabilité qu'un oiseau soit un poulet** en fonction de ses caractéristiques

$$\mathbb{P}\left(\text{oiseau} = \text{poulet} \mid \text{caractéristiques} = \begin{pmatrix} x \\ y \end{pmatrix}\right) = \sigma(ax + by + c),$$

“La probabilité qu'un oiseau soit un poulet, sachant qu'il a les caractéristiques (x, y) vaut $\sigma(ax + by + c)$.”

La fonction σ , appelée **fonction logistique**, transforme n'importe quel nombre réel en une **valeur comprise entre 0 et 1**, que l'on interprète comme une probabilité : $\sigma(x) = 1/(1 + e^{-(x-x_0)/\lambda})$.



Les paramètres a, b, c contrôlent **la position et l'orientation** de la droite séparatrice entre les deux espèces. L'apprentissage machine consiste à **trouver les valeurs** de ces paramètres qui rendent la probabilité aussi proche que possible de 1 pour les poulets, et de 0 pour les pigeons.

3 Comment la machine apprend-elle ?

La machine dispose d'un **jeu de données d'entraînement**, c'est-à-dire un ensemble d'oiseaux dont on connaît à la fois les **caractéristiques** et **l'espèce**. Elle va alors tester différentes valeurs pour les paramètres a, b, c , puis **évaluer ses performances** à l'aide d'une fonction appelée **fonction de coût** :

$$J(a, b, c) = -\frac{1}{N} \sum_{i=1}^N \left[\mathbb{1}_{\text{poulet}}(i) \ln(X_i) + \mathbb{1}_{\text{pigeon}}(i) \ln(1 - X_i) \right], \quad \text{avec } X_i = \sigma(ax_i + by_i + c)$$

- Si un oiseau est **mal classé** avec X_i proche de 0 ou de 1, **le coût augmente fortement** (le logarithme devient grand et négatif).
- Si tous les points sont **bien classés**, **le coût devient faible**.
- L'objectif de la machine est donc de **minimiser** cette fonction : apprendre, c'est chercher à réduire l'erreur globale.

4 Message à retenir

La régression logistique est un outil simple mais fondamental en apprentissage supervisé pour les problèmes de classification binaire. Elle permet non seulement de distinguer différentes espèces d'oiseaux, mais aussi de détecter les spams, de reconnaître des champignons vénéneux, ou encore d'identifier certaines cellules cancéreuses...