Intelligence artificielle

Fiche méthode

(Technologie 3e · Les objets et les systèmes techniques · L'évolution des objets et des systèmes techniques)

La démarche pour analyser une IA

1. Je regarde les données : étiquetées, sans étiquette, des récompenses, une partie cachée ? J'en déduis le type d'apprentissage.

2. Je sépare les deux phases : entraînement sur des exemples, utilisation sur des données nouvelles ; je juge le modèle sur les données de test.

3. Je place le modèle dans la fonction traiter, je nomme l'usage et je compte les faux positifs et les faux négatifs.

4. Je cherche les biais : qui manque dans les données, quel taux d'erreur pour chaque groupe ?

Le réflexe — une IA ne vaut que ce que valent ses données : je juge toujours sur des données de test, groupe par groupe, avec des taux et pas des nombres.

Cas 1 · Entraîner et tester un modèle

Réflexe — entraînement sur des exemples connus, test sur des données jamais vues ; taux de réussite = bonnes réponses ÷ tests × 100.

Exemple 1. Un modèle qui reconnaît les espèces d'arbres sur une photo d'écorce est testé sur 1 500 photos jamais vues : 1 410 sont bien reconnues. Taux de réussite ?

1 410 ÷ 1 500 × 100 = **94 %** de réussite, donc 6 % d'erreurs (90 photos).

Exemple 2. Pourquoi ne pas tester ce modèle sur ses photos d'entraînement ?

Il les a déjà vues : il peut les avoir « apprises par cœur ». Seules des photos **nouvelles** montrent s'il saura reconnaître un arbre qu'il n'a jamais vu.

Attention — diviser par le nombre de tests, pas par le nombre d'erreurs ; un score de confiance n'est pas une certitude.

Cas 2 · Reconnaître le type d'apprentissage

Réflexe — étiquettes → supervisé ; groupes sans étiquette → non supervisé ; récompenses → par renforcement ; partie cachée retrouvée → auto-supervisé.

Exemple 1. Un service de vidéo regroupe ses spectateurs selon ce qu'ils regardent, sans catégories fixées à l'avance. Quel type ?

Pas d'étiquette, des groupes à découvrir : **non supervisé**. Un humain donnera ensuite un nom à chaque groupe.

Exemple 2. Avec 1 voisin, une mesure est classée « A » ; ses 3 plus proches voisins sont A, B, B. Que répond le modèle à 3 voisins ?

2 voisins B contre 1 voisin A : **B**. La réponse dépend du nombre de voisins choisi.

Attention — des traductions faites par des humains sont des étiquettes (supervisé) ; un mot caché puis retrouvé, c'est de l'auto-supervisé.

Cas 3 · Usages, chaîne d'information, faux positifs et faux négatifs

Réflexe — je nomme l'usage, je place le modèle dans « traiter », puis : alerte à tort = faux positif ; cas réel non signalé

= faux négatif.

Exemple. Une caméra de rucher doit repérer les frelons asiatiques. Sur 300 insectes : 60 frelons, dont 57 repérés ; 9 abeilles prises pour des frelons. Compter les erreurs et le taux de réussite.

Faux négatifs : $60 - 57 = \mathbf{3}$; faux positifs : **9**. Bonnes réponses : $57 + (240 - 9) = 288$; $288 \div 300 \times 100 = \mathbf{96\ \%}$.

Attention — un modèle qui ne signale jamais rien peut avoir un bon taux de réussite : je compte aussi les faux négatifs.

Cas 4 · Repérer un biais

Réflexe — je compare la part de chaque groupe dans les données et le taux d'erreur de chaque groupe, puis je propose des données plus variées et un contrôle humain.

Exemple. Une application estime l'âge d'une personne sur une photo. Erreur moyenne : 3 ans pour les 20-40 ans, 11 ans pour les plus de 70 ans, qui forment 2 % des photos d'entraînement. Que conclure ?

Les personnes âgées, presque absentes des données, sont beaucoup moins bien traitées : c'est un **biais** de données non représentatives. Correction : ajouter des photos de personnes âgées et vérifier l'erreur groupe par groupe.

Attention — comparer des taux d'erreur, pas des nombres d'erreurs, quand les groupes n'ont pas la même taille.

Les pièges

- Calculer le taux de réussite avec les erreurs au lieu du nombre de tests.
- Juger un modèle sur ses données d'entraînement.
- Prendre un score de confiance pour une certitude.
- Confondre non supervisé et auto-supervisé ; oublier qu'une traduction humaine est une étiquette.
- Inverser faux positif (fausse alerte) et faux négatif (oubli).
- Comparer des nombres d'erreurs entre groupes de tailles différentes.

Ce document est la propriété intellectuelle de Ben, professeur particulier (benprofparticulier.fr), protégée par le Code de la propriété intellectuelle (art. L. 111-1 et L. 122-4). Il est mis à disposition pour un usage personnel de révision. Toute reproduction, diffusion, mise en ligne ou revente, totale ou partielle, sans autorisation écrite est interdite et constitue une contrefaçon (art. L. 335-2).