

Machine Learning et IA

Modèles, algorithmes et intuition

Jeff Abrahamson

octobre 2024

Bienvenue

Structure

Une matinée

- 17 octobre 2024

Structure

- Email `jeff@p27.eu`
- Github : `https://github.com/JeffAbrahamson/ML-diva-beapp`

Structure

Un peu en boucle :

- Modèles
- Algorithmes
- Exemples

Go

Le modèle le plus simple

Quel est le problème le plus simple imaginable ?

Quel est le modèle le plus simple imaginable ?

Le modèle le plus simple

Quel est le problème le plus simple imaginable ?

Quel est le modèle le plus simple imaginable ?

Un constant

Le modèle le plus simple

Quel est le problème le plus simple imaginable ?

Quel est le modèle le plus simple imaginable ?

Un constant

Par exemple ?

Le modèle le plus simple

Quel est le problème le plus simple imaginable ?

Quel est le modèle le plus simple imaginable ?

Un constant

Par exemple ?

La hauteur d'une porte. Deux cas...

Le modèle le plus simple

Quel est le problème le plus simple imaginable ?

Quel est le modèle le plus simple imaginable ?

Un constant

Par exemple ?

*La hauteur d'une porte. Deux cas : **design**, **découverte***

Le modèle le plus simple

Quel est le problème le plus simple imaginable ?

Quel est le modèle le plus simple imaginable ?

Un constant

Par exemple ?

*La hauteur d'une porte. Deux cas : **design**, **découverte***

Par exemple,

- Moyen
- Médian

Un modèle toujours simple

Mais un peu plus complexe.

Le modèle ?

Un modèle toujours simple

Mais un peu plus complexe.

Le modèle : **linéaire**

Des exemples ?

Et si ce n'est pas linéaire ?

Comment s'adapter ?

Et si ce n'est pas linéaire ?

1. Commencer avec linéaire, est-ce que ça ajoute la valeur nécessaire ?

Et si ce n'est pas linéaire ?

1. Commencer avec linéaire, est-ce que ça ajoute la valeur nécessaire ?
2. Sinon, ajouter un nouveau modèle aux résidus.

Regarder les résidus, c'est quand même toujours une bonne idée.

Et si ce n'est pas linéaire ?

1. Commencer avec linéaire, est-ce que ça ajoute la valeur nécessaire ?
2. Sinon, ajouter un nouveau modèle aux résidus.

Regarder les résidus, c'est quand même toujours une bonne idée.

Exemple A : (presse) Changement d'UX, plus de lecture ? *(penser métriques, critères)*

Et si ce n'est pas linéaire ?

1. Commencer avec linéaire, est-ce que ça ajoute la valeur nécessaire ?
2. Sinon, ajouter un nouveau modèle aux résidus.

Regarder les résidus, c'est quand même toujours une bonne idée.

Exemple A : (presse) Changement d'UX, plus de lecture ? *(penser métriques, critères)*

Exemple B : (IOT) Changement de X, plus d'autonomie ? *(penser métriques, critères)*

Vraiment pas linéaire

Distinguons entre pas linéaire parce que bombé (*linéaire peut marcher toute de même*)

— et —

rugueux, petites vagues complexes, beaucoup de minima et maxima locaux (*pas de chance avec linéaire*)

Vraiment pas linéaire

Distinguons entre pas linéaire parce que bombé (*linéaire peut marcher toute de même*)

— et —

rugueux, petites vagues complexes, beaucoup de minima et maxima locaux (*pas de chance avec linéaire*)

Rugueux peut bénéficier d'un modèle capable de trouver un séparateur plus complexe.

Vraiment pas linéaire

Distinguons entre pas linéaire parce que bombé (*linéaire peut marcher toute de même*)

— et —

rugueux, petites vagues complexes, beaucoup de minima et maxima locaux (*pas de chance avec linéaire*)

Rugueux peut bénéficier d'un modèle capable de trouver un séparateur plus complexe.

Options :

- SVM avec RBF (assez simple à mettre en oeuvre)
- Random forest (meilleure explicabilité, pour un certain esprit)
- ANN or gradient boosted trees (plus de boulot garanti, mais souvent meilleur)

Exercice

Presse

J'utilise un modèle génératif de Hugging Face afin de créer des résumé de 1–2 lignes de chaque article. J'affiche le résumé en haut de l'article, façon chapô.

Questions :

Exercice

Presse

J'utilise un modèle génératif de Hugging Face afin de créer des résumé de 1–2 lignes de chaque article. J'affiche le résumé en haut de l'article, façon chapô.

Questions :

- Est-ce que les lecteurs lisent plus d'articles ?

Exercice

Presse

J'utilise un modèle génératif de Hugging Face afin de créer des résumé de 1–2 lignes de chaque article. J'affiche le résumé en haut de l'article, façon chapô.

Questions :

- Est-ce que les lecteurs lisent plus d'articles ?
- Est-ce que les lecteurs cliquent plus de publicités ?

Exercice

Presse

J'utilise un modèle génératif de Hugging Face afin de créer des résumé de 1–2 lignes de chaque article. J'affiche le résumé en haut de l'article, façon chapô.

Questions :

- Est-ce que les lecteurs lisent plus d'articles ?
- Est-ce que les lecteurs cliquent plus de publicités ?
- Est-ce qu'un lecteur qui clique deux articles aura plus de chance d'en cliquer 5 ?

Exercice

Presse

J'utilise un modèle génératif de Hugging Face afin de créer des résumé de 1–2 lignes de chaque article. J'affiche le résumé en haut de l'article, façon chapô.

Questions :

- Est-ce que les lecteurs lisent plus d'articles ?
- Est-ce que les lecteurs cliquent plus de publicités ?
- Est-ce qu'un lecteur qui clique deux articles aura plus de chance d'en cliquer 5 ?
- Est-ce que le longueur de l'abstrait change la réponse ?

Exercice

E-commerce

Je teste un nouveau système de recommandation.

Questions :

Exercice

E-commerce

Je teste un nouveau système de recommandation.

Questions :

- Est-ce que le changement augmente le CTR ?

Exercice

E-commerce

Je teste un nouveau système de recommandation.

Questions :

- Est-ce que le changement augmente le CTR ?
- Est-ce que le changement augmente revenu ?

Exercice

E-commerce

Je teste un nouveau système de recommandation.

Questions :

- Est-ce que le changement augmente le CTR ?
- Est-ce que le changement augmente revenu ?
- Est-ce que le changement augmente le taux de retour ?

Exercice

Détection de fraude

Exercice

Détection de fraude

Définir le problème, quel surface ou frontière de décision pensez-vous avoir créer ?

Exercice

Détection de fraude

Définir le problème, quel surface ou frontière de décision pensez-vous avoir créer ?

Quels algorithmes vous semblent pertinents (et pourquoi) ?

Exercice

Détection de fraude

Définir le problème, quel surface ou frontière de décision pensez-vous avoir créer ?

Quels algorithmes vous semblent pertinents (et pourquoi) ?

Quels problèmes sont pertinents à considérer une fois que vous commencez à travailler ?

Exercice

Vos problèmes