

Entraîner et utiliser un réseau de neurones artificiels en classe

Démystifier l'Intelligence Artificielle (IA) pour les élèves et les enseignants.

Nicholas Wolff – Gymnase de Beaulieu

Biljana Petreska von Ritter-Zahony – HEP Vaud

Contexte

- Nouvelle discipline obligatoire informatique au gymnase
- En école de maturité et en école de culture générale
- Elèves qui utilisent régulièrement des IA, en particulier des LLM

Pourquoi cet outil ?

- **Le problème** : L'opacité des systèmes d'IA (la "boîte noire").
- **La solution** : Un simulateur en ligne interactif qui réduit les prérequis mathématiques au minimum.
- **Approche** : Une "boîte en verre" où l'on peut observer et manipuler l'activité interne du réseau.

Approche didactique

Il s'agit ici d'introduire l'intelligence artificielle à l'école en tant qu'objet d'apprentissage et de sortir les élèves de leur rôle d'utilisateur, pour les rendre un petit peu concepteurs de réseaux de neurones. Nous proposons d'aborder l'IA à travers une perspective technique qui ouvre la « boîte noire » pour en dévoiler les rouages et difficultés d'obtenir de bons résultats.

Approche didactique

- Se concentrer sur les principes computationnels de l'intelligence artificielle : comment elle apprend, comment on vérifie que son fonctionnement est correct, comment on interprète ses résultats
- Ne pas mettre l'accent sur la partie mathématique, afin de pouvoir s'adresser au plus grand nombre d'élèves
- En partant d'un système qui n'est pas entraîné et en arrivant à un système qui fonctionne, les élèves interagissent avec les différents éléments du réseau de neurones pour se construire une intuition visuelle de son fonctionnement et observer ses limites.

Objectifs pédagogiques

Les élèves doivent comprendre :

- Que les modèles ont besoin de beaucoup de données d'entraînement avec une grande variabilité dans ces données.
- Qu'une fois entraînés les modèles doivent être validés par un jeu de données de test qui n'ont pas été utilisées pour l'entraînement
- Que le résultat d'une IA basée sur l'apprentissage automatique ne peut pas être expliqué ni justifié

Les quatre étapes du dispositif

Etape 1 – Un neurone artificiel

Un seul neurone est capable de produire une valeur en combinant des valeurs provenant d'autres neurones en amont.

Les paramètres de cette combinaison sont les poids des connexions entre neurones.

Les élèves entraînent ce neurone à calculer une température apparente, à partir de trois valeurs d'entrée.

Étape 2 – Reconnaissance des chiffres 0 et 1

- Activité : Les élèves dessinent des chiffres pour entraîner le modèle.
- Observation : Plus il y a d'exemples, plus les prédictions deviennent fiables.
- Concept clé : L'importance centrale des données dans l'apprentissage automatique.

La notion de Généralisation

- Expérience : Tester le réseau de ses camarades avec sa propre écriture.
- Résultat : Les performances chutent si les données diffèrent de l'entraînement.
- Leçon : Ce qui n'a pas été appris ne peut pas être reconnu (limites de la généralisation).

Étape 3 – Vers la complexité (0 à 9)

- Données : Utilisation du jeu de données standard MNIST.
- Architecture : Introduction de deux couches cachées pour gérer 10 chiffres.
- Visualisation : Observation en temps réel de l'évolution des "poids" du réseau.

Paramètres de l'Apprentissage

- Variables manipulables :
 - Le taux d'apprentissage (vitesse vs stabilité).
 - La taille du jeu de données (volume nécessaire pour la pertinence).
- Analogie : Illustrer le travail empirique des ingénieurs (essais et erreurs).

Analyse des Erreurs et Biais

- Matrice de confusion : Identifier quels chiffres (comme le 4 ou le 9) sont souvent confondus.
- Probabilités : Comprendre que l'IA n'est pas déterministe mais probabiliste.
- Réflexion : Discussion sur les biais et la représentativité des données.

Etape 4 : Inexplicabilité et Mode

- Volet ludique : Test sur des données de mode.
- Constat : Mise en évidence du caractère "opaque" des décisions prises par l'IA.
- Éthique : Questionner la "confiance aveugle" accordée aux résultats.

Conclusion et Ressources

- Efficacité : Validé par des tests sur le terrain (élèves et enseignants).
- Accès : <https://nn.nwolff.info/>
- Contact : Nicholas Wolff (nicholas.wolff@eduvaud.ch).