

Nombre de participant·e·s maximal pour l'atelier	15	Nombre de sessions de 45mn que vous pouvez réaliser	2
Matériel nécessaire (vidéo-projection, grande salle, autre ...)	Vidéo-projection et connexion Internet. Pour les participant·e·s : un ordinateur portable avec une connexion à Internet.		

Simulateur didactique d'un réseau de neurones artificiels

Nicholas Wolff, Biljana Petreska von Ritter-Zahony

Haute École Pédagogique Vaud, 1007 Lausanne, Suisse

RÉSUMÉ Cet atelier présente un simulateur de réseaux de neurones en ligne, ainsi qu'une proposition de séquence pédagogique en intelligence artificielle (IA) conçue autour de cet outil. L'objectif est de démystifier le fonctionnement de ces systèmes auprès de publics lycéens, ainsi que dans le cadre de formations d'enseignant·e·s. Les prérequis en mathématiques et en informatique ont été réduits au minimum. Le réseau de neurones artificiels mis en œuvre repose sur une tâche classique de l'intelligence artificielle : la reconnaissance visuelle. Un dispositif didactique interactif permet à chaque élève, via une interface simple, de paramétrer le modèle et d'expérimenter les différentes phases d'apprentissage et de test. Les élèves manipulent successivement la collecte des données, l'entraînement et la validation du modèle. En permettant d'observer directement comment un système d'intelligence artificielle apprend et commet des erreurs, cette activité propose une introduction concrète et accessible aux réseaux de neurones et ouvre une discussion sur les enjeux sociétaux qui y sont associés.

MOTS-CLÉS • Intelligence artificielle – Apprentissage automatique – Réseaux de neurones – Fiabilité – Opacité – Enjeux sociétaux

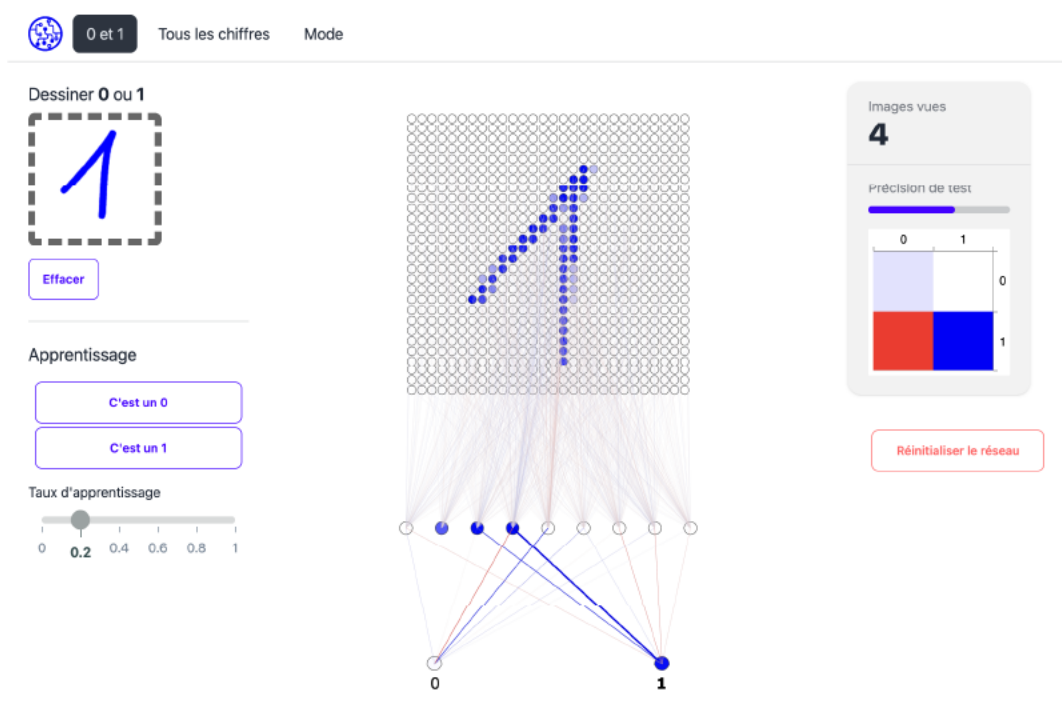


Fig. 1. Volet « 0 et 1 » du simulateur. Les élèves entraînent un réseau de neurones artificiels à reconnaître les chiffres 0 et 1 à partir d'exemples qu'ils fournissent.

1. Déroulement de l'atelier

L'atelier propose la découverte d'un simulateur de réseau de neurones artificiels (architecture à la base de nombreux systems d'intelligence artificielle (Le Cun et al., 2015)). Les participant·e·s sont amené·e·s à expérimenter une séquence pédagogique, dans le rôle de l'élève puis dans celui de l'enseignant·e, en identifiant les objectifs d'apprentissage visés ainsi que les difficultés rencontrées par les élèves.

2. Description de la ressource / activité / projet présenté dans l'atelier

Le simulateur de réseaux de neurones est structuré autour de trois volets progressifs. Dans un premier temps, les élèves entraînent un réseau de neurones à reconnaître les chiffres 0 et 1 (voir Fig. 1). Ils fournissent des exemples de ces deux chiffres et observent que, à mesure que le nombre d'exemples augmente, les prédictions du modèle deviennent progressivement plus fiables. Cette étape permet de mettre en évidence le rôle central des données dans l'apprentissage automatique.

Les élèves peuvent ensuite tester les réseaux entraînés par leurs pairs en utilisant leur propre écriture. Ils constatent alors que les performances du modèle sont généralement moins bonnes lorsque les exemples fournis diffèrent de ceux utilisés lors de l'entraînement. Cette situation permet d'introduire la notion de données d'apprentissage et d'amorcer une discussion sur les limites de la généralisation des modèles, en montrant que ce qui n'a pas été appris ne peut être



Fig. 2. Volet « Tous les chiffres » du simulateur. Les élèves manipulent le taux d'apprentissage pour entraîner un réseau de neurones artificiels à reconnaître les chiffres 0 à 9 à partir d'un jeu de données. L'interface permet d'observer en temps réel l'évolution des poids au cours de l'apprentissage et met en évidence la nécessité de disposer de jeux de données volumineux. La visualisation montre également que, même lorsque le réseau est globalement performant, certains chiffres demeurent plus difficiles à reconnaître que d'autres.

correctement reconnu. Elle ouvre également la voie à une réflexion sur les biais présents dans les jeux de données, en mettant en évidence l'influence de la diversité et de la représentativité des exemples sur les performances du modèle.

Du point de vue du fonctionnement du réseau de neurones, les élèves observent qu'un chiffre peut être représenté sous la forme de neurones d'entrée, dont l'activation correspond à la valeur de gris de chaque pixel de l'image. La prédiction est modélisée par deux neurones de sortie, associés aux chiffres 0 et 1, où le chiffre prédit correspond au neurone dont l'activation est la plus élevée.

Dans un deuxième temps, le réseau de neurones est entraîné à reconnaître l'ensemble des chiffres à partir du jeu de données MNIST (Le Cun et al., 1998). Pour reconnaître les chiffres de 0 à 9, l'architecture du modèle est complexifiée par l'introduction de deux couches cachées, conduisant à l'émergence de représentations intermédiaires (voir Fig. 2). Le dispositif est interactif : les élèves manipulent le taux d'apprentissage et la taille du jeu de données d'apprentissage, et observent leur influence sur la vitesse, la stabilité et l'efficacité de l'apprentissage. Cette activité permet d'illustrer concrètement le travail des ingénieurs et des chercheurs, fondé sur des essais successifs et des ajustements empiriques des paramètres, ainsi que la nécessité de disposer d'un volume important de données pour produire des résultats pertinents.

Les élèves peuvent ainsi suivre l'évolution du modèle au cours de l'entraînement et observer l'adaptation des poids, qui se modifient afin de produire des prédictions de plus en plus conformes aux résultats attendus. Même à l'issue du processus d'apprentissage, le système d'IA

continue de produire certaines erreurs sur le jeu de données de test, identifiables à l'aide d'une matrice de confusion. Il est alors important que les élèves prennent conscience que, contrairement aux systèmes informatiques classiques, ces systèmes présentent des comportements probabilistes, ce qui appelle à questionner les résultats produits et à moduler la confiance que l'on peut leur accorder.

Pour finir, un dernier volet, plus ludique, permet de tester un réseau de neurones entraîné sur des données de mode. Ce volet met en évidence l'inexplicabilité des décisions prises par les réseaux de neurones. L'usage de réseaux de neurones artificiels, intrinsèquement opaques, implique ainsi inévitablement une part de confiance aveugle accordée aux résultats produits.

3. Expérimentations réalisées ou envisagées

Le simulateur de réseau de neurones artificiels a été conçu dans une visée de démystification du fonctionnement des systèmes d'intelligence artificielle, largement répandus dans la société. L'objectif est d'ouvrir la « boîte noire » afin de rendre observable et manipulable l'activité interne du réseau, en particulier le rôle des neurones artificiels et l'évolution des poids au cours de l'apprentissage. Le simulateur se présente ainsi comme une boîte en verre, offrant un support à l'exploration des principes fondamentaux de l'apprentissage automatique.

Le simulateur a été utilisé à maintes reprises pour aborder les réseaux de neurones artificiels au lycée et dans des formations du corps enseignant. Deux questionnaires proposés à une classe de 17 élèves de lycée ont permis de mesurer l'impact de la séquence sur leurs connaissances et leurs attitudes face à l'intelligence artificielle, confirmant l'efficacité du dispositif.

4. Liens avec la recherche (si pertinent)

D'autres simulateurs de réseaux de neurones artificiels existent, tels que *A Neural Network Playground* (Smilkov et al., 2017) ou *Neuron Sandbox* (Touretzky et al., 2025), mais ils sont soit peu accessibles à des élèves novices en mathématiques ou en informatique, soit trop simplistes, notamment lorsque les poids du réseau restent fixes. Le simulateur proposé a donc été volontairement simplifié, dans une logique de réduction de la charge cognitive, tout en conservant uniquement deux paramètres soigneusement choisis pour agir sur le processus d'apprentissage. Ce dispositif constitue un outil didactique pour accompagner l'enseignement de l'IA, en favorisant une approche exploratoire et réflexive qui soutient la construction de modèles mentaux pertinents de l'apprentissage automatique, nécessaires à la compréhension du fonctionnement de ces systèmes et de leurs enjeux sociétaux.

5. Liens vers les ressources et point de contact

- <https://nn.nwolff.info/>
- nicholas.wolff@eduvaud.ch

6. Références

Le Cun, Y., Bengio, Y., Hinton, G. (2015). *Deep learning*. Nature, vol. 521, pp. 436–444

Le Cun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). *Gradient-based learning applied to document recognition*. *Proceedings of the IEEE*, vol. 86, pp. 2278–2324

Smilkov, D., Carter, S., Sculley, D., Viégas, F. B., Wattenberg, M. (2017). *Direct-Manipulation Visualization of Deep Networks*. CoRR abs/1708.03788
Simulateur accessible sur : <https://playground.tensorflow.org/>

Touretzky, D., Gardner-McCune, C., Hanna, W., Chen, A., Pawar, N. (2025). *Learning to Think Like a Neuron in Middle School*. In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 39, 28, pp. 29212–29219
Simulateur accessible sur : <https://www.cs.cmu.edu/~dst/NeuronSandbox/>