



HAL
open science

L'Humain dans l'Apprentissage Automatique Interactif: aperçu de l'état de l'art

Kévin Delcourt, Sylvie Trouilhet, Françoise Adreit, Jean-Paul Arcangeli

► To cite this version:

Kévin Delcourt, Sylvie Trouilhet, Françoise Adreit, Jean-Paul Arcangeli. L'Humain dans l'Apprentissage Automatique Interactif: aperçu de l'état de l'art. Conférence Nationale en Intelligence Artificielle (CNIA 2022), AFIA, Jun 2022, Saint-Etienne, France. pp.43-50. hal-03765539v1

HAL Id: hal-03765539

<https://hal.science/hal-03765539v1>

Submitted on 18 Oct 2022 (v1), last revised 31 Aug 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'Humain dans l'Apprentissage Automatique Interactif : aperçu de l'état de l'art

Kevin Delcourt, Jean-Paul Arcangeli, Sylvie Trouilhet, Françoise Adreit

Université de Toulouse, IRIT, UT3, UT2J, Toulouse, France

{Prenom.Nom}@irit.fr

Résumé

Les systèmes d'Apprentissage Automatique Interactif, ou IML (Interactive Machine Learning), font coopérer un utilisateur humain et une machine apprenante afin d'accomplir les tâches souhaitées par l'humain. Nous proposons dans cet article un aperçu de l'état de l'art de l'intégration de l'humain dans ces systèmes. Pour cela nous réalisons une étude de la littérature récente sur les solutions d'apprentissage automatique avec l'humain dans la boucle, et identifions les moyens mis en oeuvre pour prendre en compte et assister l'humain.

Mots-clés

Apprentissage Automatique, Apprentissage Interactif, Humain dans la boucle, Apprentissage Supervisé, Apprentissage par Renforcement, Interaction Humain-IA, Système Multi-Agent

Abstract

Interactive Machine Learning (IML) systems involve the cooperation of a human user and a machine learner in order to accomplish the tasks desired by the human. We propose in this paper an overview of the state of the art of human integration in these systems. To do so, we conduct a survey of recent literature on machine learning solutions with the human in the loop, and identify the means implemented to take into account and assist the human.

Keywords

Machine Learning, Interactive Learning, Human-in-the-loop, Supervised Learning, Reinforcement Learning, Human-AI Interaction, Multi-Agent System

1 Contexte

Dans le cas de problèmes complexes pour lesquels la programmation d'une solution n'est pas faisable à des coûts raisonnables, l'apprentissage automatique est la méthode de référence pour construire une solution. Pour cela, il est nécessaire de faire participer un expert en apprentissage au processus de développement de ces solutions. Or, d'après Amershi et al. [3], la demande en applications d'apprentissage automatique est bien supérieure à l'offre en compétences dans ce domaine.

Pour pallier ce manque, plusieurs approches proposent de

permettre à des utilisateurs humains souvent sans connaissances en apprentissage automatique de répondre par eux-mêmes à leurs besoins. Ces approches sont regroupées sous le terme d'Apprentissage Automatique Interactif, ou IML (*Interactive Machine Learning*). Elles permettent par ailleurs d'alléger le travail d'experts en apprentissage.

1.1 Apprentissage Automatique Interactif

Définition. Faills et Olsen Jr [16] définissent l'IML comme le domaine de l'apprentissage automatique dans lequel tout ou une partie des données exploitées par l'apprentissage provient des interactions avec un utilisateur humain du système : un humain, le plus souvent dépourvu de compétences en apprentissage automatique, interagit avec un système d'IML pour produire une solution répondant à ses besoins.

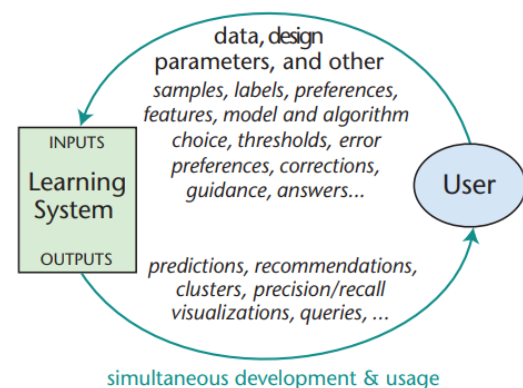


FIGURE 1 – Processus d'un système d'IML, extraite de [3]

Processus itératif. Le fonctionnement généralisé d'un système d'IML, décrit dans la figure 1, est itératif : l'humain fournit les divers paramètres, préférences ou plus généralement toute donnée demandée par l'apprenant (*i.e.* la machine) pour son fonctionnement. Ce dernier présente en retour le résultat de son apprentissage à l'humain, par exemple des recommandations ou des prédictions. L'humain doit alors déterminer si les résultats présentés par l'apprenant sont satisfaisants par rapport à l'objectif à atteindre. Le processus se répète jusqu'à ce que l'humain soit satisfait.

Applications. Les applications d’IML permettent par exemple de produire un classifieur d’images [7] ou de texte [24], d’entraîner un robot à réaliser des tâches de maintenance [8] ou de proposer des recommandations sur la base des actions de l’utilisateur [1].

1.2 Composition Logicielle Opportuniste

Nous avons été amenés à nous intéresser à l’IML par le biais de notre problématique de fond : la composition logicielle opportuniste dans le cadre de l’intelligence ambiante.

Intelligence Ambiante. L’intelligence ambiante a pour objectif d’offrir à un utilisateur humain un environnement physique et logiciel personnalisé et adapté à sa situation et à ses besoins [25]. Les principales difficultés à surmonter proviennent de la dynamique, de l’ouverture et de l’imprévisibilité des systèmes ambiants, induites par la mobilité des appareils et des utilisateurs. Le nombre de composants présents dans l’environnement est une autre source de complexité. Pour proposer des applications utiles et utilisables, les solutions doivent donc être capables de prendre en compte le contexte opérationnel et ses changements, notamment les préférences et les besoins de l’utilisateur ainsi que leurs évolutions.

Composition Opportuniste. Notre projet de recherche vise à concevoir et à développer une solution qui intègre l’utilisateur humain dans la boucle et fait émerger de nouvelles fonctionnalités dans un environnement ambiant. Il s’agit de construire, pour un humain, les bonnes applications au bon moment sans qu’il doive explicitement les demander. Notre solution utilise un apprentissage par renforcement [28], distribué au sein d’un système multi-agent (SMA), qui tire parti de retours de l’utilisateur pour construire une connaissance sur ses choix préférentiels en fonction des composants présents dans l’environnement.

Prototype. Un modèle de composition opportuniste a été conçu et un prototype logiciel a été développé [32, 22]. Une démonstration de ce dernier sur la base d’un cas d’utilisation est présentée dans [12]. Le fonctionnement itératif du moteur de composition opportuniste OCE (*Opportunistic Composition Engine*) est décrit dans la figure 2 : OCE sonde l’environnement de l’humain ; son SMA assemble ensuite une application qui est proposée à l’humain pour validation ; enfin, l’interface de visualisation de l’assemblage, ICE (*Interactive Control Environment*) [22], permet à l’humain d’accepter, de modifier ou de rejeter la proposition d’OCE. Cette action constitue le retour d’information qui est la source d’apprentissage pour les agents du SMA d’OCE.

1.3 Problématique

Notre solution de composition logicielle opportuniste est un exemple de système d’Apprentissage Automatique Interactif. Nous y retrouvons en effet un processus itératif avec, dans la boucle, un humain et une machine apprenante. L’humain occupe ici une place essentielle : en interaction avec le système apprenant, il est la principale source de l’apprentissage et c’est pour lui que l’apprenant opère.

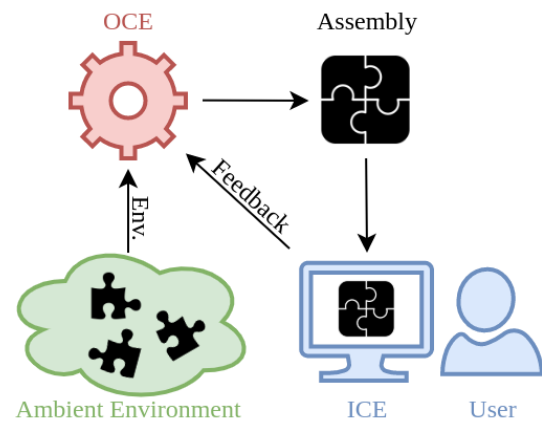


FIGURE 2 – Composition Logicielle Opportuniste

Alors qu’une bonne intégration de l’humain est indispensable à une solution d’IML de qualité [3], la question de la place de l’humain dans notre système n’a pas été approfondie jusqu’à présent.

Plusieurs questions se posent quant à l’intégration de l’humain dans la composition logicielle opportuniste :

- Comment informer l’utilisateur de manière appropriée par rapport à ses capacités sur les applications que fait émerger le moteur de composition ?
- Comment faire contribuer l’utilisateur au système de manière à améliorer l’apprentissage ?
- Comment faire en sorte que ces interactions ne sollicitent pas trop l’utilisateur ?

Des réponses partielles existent dans OCE, comme les options de visualisation de ICE. Notre objectif est de proposer d’autres éléments de réponse. Pour cela, les questions de recherche définies dans la section suivante vont guider notre lecture des contributions récentes du domaine de l’IML.

1.4 Questions de recherche

Nous cherchons à répondre aux questions suivantes :

QR1-Humain. Quels sont le rôle et les responsabilités de l’utilisateur dans la boucle ?

- **QR11-Tâches.** Quelles tâches l’humain doit-il accomplir ? Pour accomplir ces tâches, quelle est la part de compétences en apprentissage automatique ?
- **QR12-Charge.** Quelle charge est imposée à l’humain ? Quel niveau d’engagement ou d’implication est attendu ?

QR2-Apprenant. Comment l’humain est-il pris en compte par la machine apprenante ?

- **QR21-Information.** Quelles sont les informations fournies par l’humain ? Comment et quand sont-elles considérées dans le processus ?
- **QR22-Assistance.** Comment le système accompagne-t-il et assiste-t-il l’humain ?

La section 2 présente notre lecture de la littérature récente par rapport à ces questions. La section 3 synthétise et analyse les réponses à ces questions.

2 Relations entre humain et machine apprenante

Afin de répondre aux questions de recherche élicitées dans la section 1.4, nous avons identifié des contributions scientifiques pertinentes dans la littérature récente du domaine l'IML. Sur la base de 32 articles traitant de différents sujets relatifs à des applications d'IML avec l'humain dans la boucle, nous avons sélectionné 12 contributions pertinentes par rapport à la question de la co-construction humain-machine d'applications. D'autres travaux d'état de l'art sur l'IML ou des domaines connexes ont été publiés [9, 10, 14, 20, 23, 33] mais ils ne ciblent pas la place de l'humain dans la boucle.

Dans cette section, nous présentons les contributions sélectionnées. Nous résumons chacune d'entre elles puis répondons aux quatre questions de recherche. Nous présentons d'abord les travaux d'apprentissage supervisé interactif, qui est à notre connaissance le plus étudié dans la littérature, puis les travaux d'apprentissage par renforcement interactif, plus proches de notre problématique. Nous terminons par les contributions traitant d'autres types d'apprentissage.

2.1 Apprentissage Supervisé

Les travaux sélectionnés sur l'apprentissage supervisé interactif concernent la production par l'humain de classificateurs d'image, de texte, ou de mouvements. La première contribution présentée concerne une approche particulière d'IML, la deuxième une étude, et les suivantes des outils.

2.1.1 Enseignement Automatique Interactif

Défini par Ramos et al. [24], l'Enseignement Automatique Interactif (*Interactive Machine Teaching*) est une approche d'IML dans laquelle l'humain joue le rôle d'un professeur, et doit enseigner une tâche à la machine. La notion d'enseignement regroupe ici le choix des informations à la source de l'apprentissage, et l'évaluation des performances de l'apprenant. L'hypothèse de fond est que l'humain acquiert plus facilement des compétences en pédagogie qu'en apprentissage automatique, ces compétences étant plus répandues dans le grand public [30].

Simard et al. [27] font un parallèle entre l'Enseignement Automatique Interactif et la programmation en général. Les deux activités partagent par exemple le fait de produire un artefact (un modèle appris ou un programme) répondant aux besoins d'un ou plusieurs humains. Cette comparaison amène les auteurs à penser que, à l'instar de la programmation qui a bénéficié d'outils et de langages de haut niveau, l'IML a besoin de ses propres outils et abstractions pour faciliter le travail des humains.

Les concepts et processus étudiés dans le cadre de l'Enseignement Automatique Interactif sont applicables à tout paradigme d'apprentissage. Néanmoins, les auteurs présentent une application de démonstration, appelée PICL [24] (anciennement MATE [30]), qui applique les principes de l'Apprentissage Automatique Interactif à l'apprentissage supervisé en permettant à des utilisateurs d'enseigner des tâches de classification de texte. Nous répondons aux ques-

tions de recherche pour cette application.

QR11-Tâches. L'humain, jouant ici le rôle d'un enseignant, doit planifier un curriculum, puis le mettre à jour en fonction des résultats de l'apprenant. Un curriculum désigne les données (exemples, étiquettes) exploitées par l'apprenant. Les compétences demandées sont plus de l'ordre de la pédagogie que de l'apprentissage automatique.

QR12-Charge. Demander à l'humain de déterminer quels sont les exemples les plus pertinents ou quels concepts enseigner l'implique fortement dans le processus. De plus, il doit être capable de comprendre si les résultats de l'apprenant sont satisfaisants ou non.

QR21-Information. Les labels et les concepts fournis par l'humain sont la source de l'apprentissage supervisé.

QR22-Assistance. L'interface des outils présentés [24] assiste l'humain en lui permettant de visualiser efficacement les résultats de l'apprenant. Une étude du comportement d'experts en apprentissage supervisé [30] a également permis d'identifier de bonnes pratiques d'enseignement automatique, implantées sous la forme de notifications.

2.1.2 Étude sur la confiance dans l'IML

L'étude menée par Honeycutt et al. [19] porte sur la question de la confiance de l'humain dans les applications d'apprentissage automatique. Elle se base sur des résultats en psychologie [29] qui montrent que la confiance d'un individu envers un groupe décisionnel humain augmente si un avis émis par l'individu, est pris en compte par le groupe. Inversement, l'individu aura moins confiance si son avis est ignoré. L'objectif de l'étude est de retrouver ces résultats en remplaçant le groupe décisionnel humain par un apprenant automatique, ici une application d'apprentissage supervisé de reconnaissance de visages dans des images.

L'expérimentation menée en ligne a mesuré la confiance d'utilisateurs humains envers cette application avec ou sans interaction d'une part, et avec une performance croissante, constante ou décroissante de l'apprenant d'autre part. Concrètement, la moitié des participants devait corriger les erreurs de l'apprenant, et la fréquence de ces erreurs était croissante pour un tiers, constante pour un autre tiers et décroissante pour le dernier.

Les résultats montrent qu'en général le groupe en interaction a moins confiance dans le système que le groupe qui n'est pas en interaction. L'explication avancée par les auteurs est que le groupe en interaction a passé plus de temps focalisé sur les erreurs du système pour les corriger.

Nous répondons aux questions de recherche pour l'application de reconnaissance de visages exploitée dans le cadre de l'expérimentation.

QR11-Tâches. Les compétences attendues pour les utilisateurs sont uniquement relatives à la tâche concernée : la reconnaissance de visage.

QR12-Charge. En dehors de la tâche de vérification des sorties de l'apprenant, l'utilisateur est faiblement impliqué.

QR21-Information. L'application simulée prend en compte les retours des utilisateurs sur ses erreurs afin d'affiner son modèle.

QR22-Assistance. Les utilisateurs sont volontairement peu accompagnés afin de pouvoir mesurer leur appréciation subjective de la performance de l'application, qui est une mesure proportionnelle à la confiance ressentie [31].

2.1.3 Classification d'image, son et postures

Carney et al. [7] présentent l'outil Teachable Machine de Google¹ permettant à des utilisateurs finaux de construire un classifieur d'images, de sons, ou de postures du corps humain. L'outil s'adresse à des utilisateurs novices en ML, avec possiblement des compétences en codage, puisque les modèles peuvent être exportés dans le langage JavaScript. Afin de faciliter la tâche des utilisateurs, un transfert d'apprentissage a été mis en place : l'entraînement d'un nouveau modèle par l'humain est fait sur la base d'un modèle de base polyvalent, entraîné en amont par les développeurs de l'application.

QR11-Tâches. Ici l'humain a la charge de sélectionner les concepts à enseigner, et de recueillir les différents exemples (*i.e.* les images) associés. Il doit observer et juger lui-même les résultats de l'apprenant, donc fournir les données de test.

QR12-Charge. Il intervient à chaque étape de la production du modèle appris, du choix des données d'entraînement à l'évaluation de l'apprenant. Cela l'implique fortement.

QR21-Information. Les entrées fournies par l'humain, sous la forme d'images étiquetées, permettent de personnaliser le modèle de base fourni par l'application en fonction des besoins. Il est également possible de modifier les paramètres avancés de l'apprenant.

QR22-Assistance. L'interface guide l'utilisateur dans les différentes étapes de production du modèle. De plus, elle cache dans un premier temps les réglages avancés de l'apprenant, afin que l'investissement pour produire un modèle soit le plus faible possible.

2.1.4 Reconnaissance de formes

Fails et Olsen Jr [16] proposent à un humain de produire un système de reconnaissance de formes, au moyen d'une application d'IML appelée Crayons. Via une interface, l'humain colorie (*i.e.* étiquette) des portions de l'image qu'il souhaite que l'apprenant classifie. Ce dernier propose ensuite une classification de l'ensemble de l'image à l'humain, qui peut affiner son étiquetage pour corriger les éventuelles erreurs de Crayons.

Les auteurs discutent les différentes exigences liées à l'IML, en particulier le fait que l'apprentissage doit se faire le plus rapidement possible pour ne pas décourager l'humain. Par rapport à ces exigences, ils identifient les arbres de décision comme étant la meilleure approche, étant donné un temps d'apprentissage plus court malgré de moins bonnes performances que d'autres méthodes comme les réseaux de neurones. Notons cependant que cette conclusion n'est peut-être plus d'actualité étant donné l'ancienneté de cet article (2003).

QR11-Tâches. L'humain doit savoir déterminer sur l'image les classes à identifier (par exemple la peau) et en colorier une partie.

1. <https://teachablemachine.withgoogle.com/>

QR12-Charge. L'humain n'a pas à colorier avec exactitude les zones à classifier, ce qui évite donc de le surcharger mentalement. Il doit cependant affiner cet étiquetage jusqu'à ce que les résultats lui conviennent.

QR21-Information. L'image fournie par l'humain et les étiquettes qu'il appose en coloriant sont la source de l'apprentissage supervisé.

QR22-Assistance. L'étiquetage par coloriage permet de simplifier la tâche de l'humain. En dehors de cela, l'humain n'est pas particulièrement accompagné.

2.1.5 Classification d'images médicales

Berg et al. [5] présentent Ilastik, un outil d'Apprentissage Automatique Interactif pour la classification d'images dans le domaine médical. Ilastik permet à un expert du domaine médical de produire un classifieur d'images, de manière similaire à Crayons (section 2.1.4). Il est cependant plus complet, prenant en compte des données jusqu'à 5 dimensions (3 pour l'espace, 1 pour le temps, et 1 pour la multiplicité des points de vue). L'outil propose également à l'humain un choix parmi sept modes de classification d'images, qui impliquent différentes interactions avec l'humain. Par exemple, la manière d'étiqueter les exemples fournis par l'humain change en fonction de l'algorithme choisi (coloriage, clic sur des formes à détecter, etc.).

QR11-Tâches. Outre des compétences dans le domaine médical, l'humain doit savoir choisir, en fonction de son objectif, quel mode de classification utiliser. Il doit ensuite fournir et annoter des exemples représentatifs.

QR12-Charge. La multitude des modes de classification proposés, qui impliquent chacun des actions différentes de la part de l'humain, font que ce dernier doit être fortement investi dans l'application afin de la maîtriser.

QR21-Information. Les exemples annotés fournis par l'utilisateur sont la seule source d'entraînement.

QR22-Assistance. Si ce n'est la documentation disponible en ligne, l'humain est peu accompagné pour accomplir sa tâche. Une bibliothèque de modèles pré-entraînés est proposée, sans pouvoir toutefois modifier ces modèles pour les adapter à une tâche particulière.

2.1.6 Détection de l'activité

Flutura et al. [17] ont développé Drinkwatch, une application embarquée dans une montre connectée permettant à des humains de suivre leur consommation de boissons. Pour cela, l'application apprend les habitudes de consommation de l'humain via les capteurs de la montre.

QR11-Tâches. Tout le long de son activité, l'humain est sollicité pour valider ou invalider certaines activités détectées par la montre, c'est-à-dire qu'il collabore pour détecter les faux positifs. Il doit également signaler quand une activité n'a pas été détectée, c'est-à-dire les faux négatifs.

QR12-Charge. L'humain est sollicité uniquement quand l'application a un degré de confiance faible dans son classement, afin d'éviter de surcharger l'humain. Il reste cependant impliqué étant donné qu'il doit être vigilant aux faux négatifs et positifs de l'application.

QR21-Information. L'apprentissage se base sur les mouvements détectés par la montre et les retours de l'humain.

QR22-Assistance. Outre le fait que l'application sollicite le moins possible l'humain, des notifications sonores signalent la détection d'une activité par Drinkwatch afin de faciliter le travail de l'utilisateur.

2.2 Apprentissage par Renforcement

Les différents projets de recherche sélectionnés ici proposent des approches indépendantes de tout domaine d'application : tous permettent à un humain d'entraîner un agent (robot virtuel ou réel) à effectuer une tâche.

2.2.1 Apprendre d'humains inattentifs

Kessler Faulkner et Thomaz [21] s'intéressent au problème de l'inattention des utilisateurs humains. Pour cela, ils considèrent un cas d'utilisation où un humain doit enseigner des actions à un robot. L'humain observe le robot agir et peut récompenser positivement les actions du robot qui lui conviennent. Le robot a alors le choix entre l'exploitation d'actions qu'il sait satisfaisantes pour l'humain, et l'exploration de nouvelles actions à la récompense incertaine.

Ainsi, en faisant l'hypothèse que l'agent apprenant sache détecter l'absence d'attention de l'humain, les auteurs proposent que l'agent favorise l'exploitation quand l'humain est inattentif, et l'exploration dans le cas contraire. Plutôt que de demander l'attention de l'humain, ce qui peut conduire à une expérience utilisateur négative, le robot tire parti au maximum de l'attention disponible.

QR11-Tâches. L'humain doit surveiller les actions de l'apprenant et récompenser via une interface simple celles qui lui conviennent.

QR12-Charge. Bien que la surveillance de l'apprenant a tendance à fortement mobiliser l'humain, le fait qu'une baisse d'attention n'ait pas de conséquence néfaste pour l'apprenant est bénéfique pour l'expérience utilisateur.

QR21-Information. Les récompenses données par l'utilisateur sont celles de l'apprentissage par renforcement que l'apprenant cherche à maximiser.

QR22-Assistance. Ici, l'humain n'est pas accompagné.

2.2.2 Apprendre malgré des retours erronés

Akrour et al. [2] présentent une approche de l'apprentissage par renforcement interactif nommée programmation par feedback. Toujours dans un contexte d'humain récompensant les actions d'un apprenant, une particularité de cette approche est que l'apprenant estime le taux d'erreur de l'humain, c'est-à-dire une estimation de la probabilité que le retour donné par l'humain soit faux. Ainsi les avis de l'humain qui semblent être à l'opposé d'avis précédents sont ignorés.

QR11-Tâches. Après chaque séquence d'actions de l'apprenant, l'humain doit comparer la dernière séquence avec la meilleure séquence parmi les précédentes, et choisir celle qui correspond le mieux à la tâche à réaliser. Il doit donc uniquement savoir évaluer le comportement de l'apprenant dans sa globalité.

QR12-Charge. L'humain doit étudier chaque séquence d'actions de l'apprenant, ce qui demande une implication forte.

QR21-Information. Les comparaisons produites par l'humain permettent à l'apprenant d'estimer une fonction de récompense qu'il essaye de maximiser par ses actions.

QR22-Assistance. Ici, l'humain est peu accompagné dans son évaluation des séquences d'actions.

2.2.3 Apprentissage profond par renforcement

Christiano et al. [11] ont conçu un algorithme d'apprentissage profond par renforcement à partir de retours humains. Pour modéliser un réseau de neurones sur la base des retours de l'humain, l'apprenant soumet à l'humain de courtes séquences d'actions à comparer. Cette comparaison est la source d'apprentissage d'une fonction de récompense que l'algorithme cherche en permanence à optimiser.

Les auteurs ont expérimenté leur solution sur plusieurs cas d'utilisation via la plateforme OpenAI Gym [6]. Dans les cas où les fonctions de récompenses sont connues, la solution a pu être comparée avec un algorithme d'apprentissage par renforcement classique et dans certains cas elle atteint de meilleurs résultats. Nous notons cependant qu'un seul utilisateur humain a participé aux expérimentations, ce qui limite la portée des conclusions.

QR11-Tâches. Ici l'utilisateur doit comparer deux séquences d'actions et décider de la meilleure par rapport à son objectif.

QR12-Charge. Le fonctionnement asynchrone apprenant-humain fait que moins de 1% des actions de l'apprenant sont jugées par l'humain. La charge qui lui incombe est donc réduite.

QR21-Information. Les retours de l'humain permettent de mettre à jour la fonction de récompense de l'algorithme.

QR22-Assistance. L'interface offre des raccourcis clavier facilitant l'interaction de l'humain.

2.3 Autres travaux

2.3.1 Apprentissage interactif pour l'optimisation

Holzinger et al. [18] présentent une expérimentation comparant des approches d'apprentissage avec et sans assistance humaine dans le cadre de résolution de problèmes du voyageur de commerce.

Sur la base d'un système multi-agent de colonie de fourmis (*Ant Colony Optimization*) [13], deux solutions, faisant participer ou non l'humain ; sont comparées. Dans l'étude présentée, l'ajout de l'humain en assistance du système multi-agent améliore les résultats de l'algorithme d'optimisation.

QR11-Tâches. L'utilisateur agit ici de manière implicite sur l'apprentissage : l'interface utilisée prend la forme d'un jeu. On lui demande donc simplement de jouer sans l'informer du système multi-agent qu'il influence.

QR12-Charge. Il est faiblement impliqué, car il n'est pas informé de l'apprentissage, ni de ses résultats.

QR21-Information. Les actions de l'humain dans le jeu ont pour effet d'aider le système multi-agent.

QR22-Assistance. Ici, l’humain n’est pas accompagné mais il n’a pas à l’être puisque sa contribution n’est pas explicite : il n’est pas informé de l’algorithme en arrière-plan.

2.3.2 Recommandation de lecture

Schnabel et al. [26] font une étude de plusieurs interfaces pour une même application de recommandation d’articles à lire. Ils s’intéressent à la visibilité qu’a un utilisateur sur l’effet de son action (*i.e.* sélectionner ou non un article à lire) sur l’apprentissage et les actions futures de l’apprenant (*i.e.* les futures recommandations). L’objectif est de mesurer l’impact de cette visibilité sur l’efficacité des utilisateurs mesurée en nombre de bonnes recommandations acceptées par l’humain en un temps donné. L’étude s’intéresse également à l’expérience utilisateur ressentie.

Pour répondre aux questions de recherche nous nous intéressons à la version de l’application préférée des utilisateurs lors de l’étude. Dans cette dernière les conséquences de la sélection d’un article à lire sont mises en valeur pour l’humain avant même que l’action soit effectuée. C’est-à-dire qu’au survol d’un article par la souris, les modifications que son ajout apporterait à la liste de des recommandations sont affichées en surbrillance.

QR11-Tâches. L’utilisateur doit sélectionner des articles à lire parmi des recommandations. Il doit juger de la pertinence de ces articles en fonction de ses goûts.

QR12-Charge. Le niveau d’implication de l’humain est faible compte tenu de la tâche à effectuer. Dans le cadre de l’expérimentation, une contrainte de temps a été instaurée, qui ajoute une certaine charge mentale aux humains.

QR21-Information. Les actions de l’utilisateur, d’ajouter un article à leur liste de lecture sont la base de la personnalisation du système de recommandation.

QR22-Assistance. La pré-visualisation de l’impact des actions de l’humain est une assistance qui a été fortement appréciée lors de l’étude. Elle a permis au groupe concerné d’être plus efficace dans le choix d’articles.

2.3.3 Lignes directrices pour tout type d’application

Amershi et al. [4] proposent 18 lignes directrices pour la conception des modalités d’interaction des applications d’Apprentissage Automatique Interactif. L’objectif est d’aider les concepteurs à éviter les expériences utilisateurs négatives dues à une interface mal conçue.

Afin de produire ces lignes directrices, les auteurs sont partis d’une étude de la littérature des 20 dernières années sur le sujet. Cette étude a fait ressortir 168 recommandations spécifiques, qui ont été regroupées, filtrées et raffinées pour arriver à 18 propositions. Ce processus de raffinement a consisté en une étude auprès de praticiens et d’experts en IHM, qui ont apprécié la pertinence de ces lignes directrices dans diverses applications d’Apprentissage Automatique Interactif disponibles dans le commerce.

Ces lignes directrices générales constituent une aide pertinente aux professionnels de l’IHM. On y retrouve par exemple le fait d’expliquer clairement les conséquences des actions de l’humain sur l’apprenant, ou inversement d’expliquer comment une décision a été prise par la machine.

QR22-Assistance. Les lignes directrices aident à mettre au point un accompagnement qui allège l’utilisateur. Étant donné l’aspect généraliste de ces recommandations, nous ne pouvons pas répondre aux autres questions de recherche pour cette contribution.

3 Analyse et conclusion

3.1 Réponses aux questions de recherche

La table 1 synthétise les niveaux de réponses aux quatre questions élicitées section 1.4. Pour chacune, nous discutons des éléments importants ou originaux trouvés dans la littérature.

Contribution	QR11	QR12	QR21	QR22
Ramos et al. [24]	+	-	+	++
Honeycutt et al. [19]	++	++	-	-
Carney et al. [7]	+	+	++	+
Failis et Olsen Jr [16]	+	+	-	-
Berg et al. [5]	-	-	++	-
Flutura et al. [17]	+	+	+	+
Kessler et al. [21]	++	+	+	-
Akrour et al. [2]	+	-	-	-
Christiano et al. [11]	+	+	-	+
Holzinger et al. [18]	++	++	-	-
Schnabel et al. [26]	+	+	+	++
Amershi et al. [4]	N/A	N/A	N/A	++

TABLE 1 – Synthèse des niveaux de réponses aux différentes questions de recherche. L’échelle "-" à "++" indique le niveau de réponse à une question. N/A = Non Applicable.

QR11-Tâches. Les applications d’IML sont variées, par conséquent les tâches demandées à l’humain le sont aussi : enseigner, superviser, reconnaître une image, choisir parmi une liste de propositions... La colonne QR11 indique en quoi ces tâches sont propres au métier et non à l’apprentissage automatique.

De manière générale, peu de compétences en apprentissage sont requises dans les contributions présentées (+ ou ++). Malgré tout, les contributions notées + demandent à l’humain de manipuler des concepts propres à l’apprentissage automatique, par exemple des étiquettes dans l’apprentissage supervisé [24, 7, 16, 17] et des séquences d’actions dans l’apprentissage par renforcement [2, 11].

Enfin, Ilastik [5] demande à l’humain une familiarité avec différents algorithmes de reconnaissance d’images. Cette tâche relève fortement du domaine de l’apprentissage.

QR12-Charge. La colonne QR12 donne une évaluation de la charge et de l’implication demandées à l’humain : elle est modérée (notée +) dans la moitié des applications. Les applications où l’humain est le moins surchargé sont celles où la tâche demandée à l’utilisateur est la plus simple : pointer les erreurs d’un apprenant [19] ou alimenter indirectement l’apprentissage [18].

QR21-Information. L’apprenant exploite de diverses manières les informations fournies par l’humain. La colonne QR21 indique le niveau de complexité de l’interven-

tion de l'humain. Les contributions notées - ne proposent qu'une seule manière pour l'humain d'interagir, comme par exemple l'avis sur les actions de l'apprenant dans [11].

À l'inverse, les contributions notées ++ proposent des interactions plus riches donnant plus de possibilités pour l'humain d'influer sur l'apprentissage machine. Il est par exemple possible dans [7] d'agir plus en profondeur dans le paramétrage de l'apprenant, ce qui nécessite plus de compétences en apprentissage de la part de l'humain. Certaines sources d'interaction implicites, comme le survol de la souris [26] ou la détection de l'attention [21] permettent également à l'humain d'agir plus indirectement sur l'apprenant.

QR22-Assistance. La colonne QR22 donne le niveau d'assistance qu'apporte l'application d'IML à l'humain : il est faible pour la majorité des applications. À l'exception des travaux de Honeycutt et al. [19] (assistance volontairement minimale) et de Holzinger et al. [18] (assistance non nécessaire), une meilleure assistance pourrait être profitable.

3.2 Positionnement et contributions

Pour terminer, nous discutons quelques principes particulièrement intéressants pour la conception d'interactions humain-IA en général. Ils constituent également pour nous des futures pistes de réflexion pour la composition logicielle opportuniste.

Utiliser des lignes directrices. Les lignes directrices proposées par Amershi et al. [4], bien que générales, permettent d'identifier des pistes pour faciliter le travail de l'humain. Par exemple, le fait de permettre à l'humain de pré-visualiser les conséquences de ses actions [26].

Entraîner l'apprenant avec des experts en apprentissage. Une approche également intéressante est d'observer des utilisateurs experts en apprentissage afin d'en tirer des recommandations, implantées par la suite sous forme de notifications [30]. Cela participe à former les humains expérimentés aux particularités d'une application avec un apprenant dans la boucle.

Opérer des transferts de connaissances. Il est aussi possible de permettre aux humains de partager leurs modèles, comme ce qui est présenté dans [5]. Récupérer un modèle existant répondant à une tâche proche de celle que l'on souhaite résoudre, puis le spécialiser (ce que proposent Carney et al. [7]) semble constituer un moyen intéressant de tirer parti d'un grand nombre d'utilisateurs. Ce principe pourrait permettre de prendre en compte l'apparition de composants jusque là inconnus dans l'environnement ambiant, ainsi que la personnalisation face à un nouvel utilisateur.

Prendre en compte les changements d'objectifs, de préférences de l'humain. Dans la majeure partie des contributions, l'humain interagit avec un apprenant automatique pour une tâche clairement définie et la possibilité qu'elle change avec le temps n'est pas envisagée.

Par exemple, dans les approches proposées dans [2] et [21], un changement d'objectif de l'humain entraînera des données possiblement contradictoires pour l'apprentissage, qui seront traitées comme des erreurs.

Dans les approches basées sur la construction de classificateurs d'images ou de texte [5, 7, 24, 16], l'humain doit construire un nouveau modèle si son objectif change, éventuellement en modifiant les données d'entraînement d'un modèle existant. Les solutions d'IML pourraient pourtant être sensibles à la dynamique pour continuer à travailler correctement en cas de changements.

Utiliser la participation implicite comme source d'apprentissage. La participation implicite [15], telle que la prise en compte des mouvements de la souris [21], peut enrichir l'apprentissage sans surcharger l'humain ou le contraindre à acquérir de nouvelles compétences (par exemple en pédagogie pour [24]).

Références

- [1] M. Aamir and M. Bhusry. Recommendation system : state of the art approach. *International Journal of Computer Applications*, 120(12), 2015.
- [2] R. Akrou, M. Schoenauer, M. Sebag, and J.-C. Souplet. Programming by feedback. In *Int. Conf. on Machine Learning*, volume 32, pages 1503–1511. JMLR.org, 2014.
- [3] S. Amershi, M. Cakmak, W. B. Knox, and T. Kulesza. Power to the people : The role of humans in interactive machine learning. *Ai Magazine*, 35(4) :105–120, 2014.
- [4] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, and K. Inkpen. Guidelines for human-AI interaction. In *Proc. of the 2019 CHI conf. on human factors in computing systems*, pages 1–13, 2019.
- [5] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, and M. Rudy. Ilastik : interactive machine learning for (bio) image analysis. *Nature Methods*, 16(12) : 1226–1232, 2019.
- [6] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. OpenAI Gym. *arXiv preprint arXiv :1606.01540*, 2016.
- [7] M. Carney, B. Webster, I. Alvarado, K. Phillips, N. Howell, J. Griffith, J. Jongejan, A. Pitaru, and A. Chen. Teachable machine : Approachable Web-based tool for exploring machine learning classification. In *Extended abstracts of the 2020 CHI Conf. on human factors in computing systems*, pages 1–8, 2020.
- [8] C. Celemin and J. Ruiz-del Solar. An interactive framework for learning continuous actions policies based on corrective feedback. *Journal of Intelligent & Robotic Systems*, 95(1) :77–97, 2019.
- [9] A. Chatzimpampas, R. M. Martins, I. Jusufi, and A. Kerren. A survey of surveys on the use of visualization for interpreting machine learning models. *Information Visualization*, 19(3) :207–233, 2020.

- [10] A. Chatzimparmpas, R. M. Martins, I. Jusufi, K. Kucher, F. Rossi, and A. Kerren. The state of the art in enhancing trust in machine learning models with the use of visualizations. In *Computer Graphics Forum*, volume 39, pages 713–756. Wiley Online Library, 2020.
- [11] P. Christiano, J. Leike, T. B. Brown, M. Martic, S. Legg, and D. Amodei. Deep reinforcement learning from human preferences. *arXiv preprint arXiv :1706.03741*, 2017.
- [12] K. Delcourt, F. Adreit, J.-P. Arcangeli, K. Hacid, S. Trouilhet, and W. Younes. Automatic and Intelligent Composition of Pervasive Applications - Demonstration. In *19th IEEE Int. Conf. on Pervasive Computing and Communications (PerCom 2021)*, Kassel (virtual), Germany, March 2021.
- [13] M. Dorigo, M. Birattari, and T. Stutzle. Ant colony optimization. *IEEE computational intelligence magazine*, 1(4) :28–39, 2006.
- [14] J. J. Dudley and P. O. Kristensson. A review of user interface design for interactive machine learning. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(2) :1–37, 2018.
- [15] C. Evers, R. Kniewel, K. Geihs, and L. Schmidt. The user in the loop : Enabling user participation for self-adaptive applications. *Future Generation Computer Systems*, 34 :110–123, May 2014.
- [16] J. A. Fails and D. R. Olsen Jr. Interactive machine learning. In *Proceedings of the 8th Int. Conf. on Intelligent user interfaces*, pages 39–45, 2003.
- [17] S. Flutura, A. Seiderer, I. Aslan, C.-T. Dang, R. Schwarz, D. Schiller, and E. André. Drinkwatch : A mobile wellbeing application based on interactive and cooperative machine learning. In *Proceedings of the 2018 Int. Conf. on Digital Health*, pages 65–74, 2018.
- [18] A. Holzinger, M. Plass, M. Kickmeier-Rust, K. Holzinger, G. C. Crişan, C.-M. Pintea, and V. Palade. Interactive machine learning : experimental evidence for the human in the algorithmic loop. *Applied Intelligence*, 49(7) :2401–2414, 2019.
- [19] D. Honeycutt, M. Nourani, and E. Ragan. Soliciting human-in-the-loop user feedback for interactive machine learning reduces user trust and impressions of model accuracy. In *Proceedings of the AAAI Conf. on Human Computation and Crowdsourcing*, volume 8, pages 63–72, 2020.
- [20] L. Jiang, S. Liu, and C. Chen. Recent research advances on interactive machine learning. *Journal of Visualization*, 22(2) :401–417, 2019.
- [21] T. A. Kessler Faulkner and A. Thomaz. Interactive Reinforcement Learning from Imperfect Teachers. In *Companion of the 2021 ACM/IEEE Int. Conf. on Human-Robot Interaction*, pages 577–579, 2021.
- [22] M. Koussaifi. *Modélisation centrée utilisateur pour la configuration logicielle en environnement ambiant*. Thèse de doctorat, Université Paul Sabatier-Toulouse III, 2020.
- [23] G. Li, R. Gomez, K. Nakamura, and B. He. Human-centered reinforcement learning : A survey. *IEEE Transactions on Human-Machine Systems*, 49(4) :337–349, 2019.
- [24] G. Ramos, C. Meek, P. Simard, J. Suh, and S. Ghorashi. Interactive machine teaching : a human-centered approach to building machine-learned models. *Human-Computer Interaction*, 35(5-6) :413–451, 2020.
- [25] F. Sadri. Ambient intelligence : A survey. *ACM Computing Surveys*, 43(4) :1–66, 2011.
- [26] T. Schnabel, S. Amershi, P. N. Bennett, P. Bailey, and T. Joachims. The Impact of More Transparent Interfaces on Behavior in Personalized Recommendation. In *Proceedings of the 43rd Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, pages 991–1000, 2020.
- [27] P. Y. Simard, S. Amershi, D. M. Chickering, A. E. Pelton, S. Ghorashi, C. Meek, G. Ramos, J. Suh, J. Verwey, and M. Wang. Machine teaching : A new paradigm for building machine learning systems. *arXiv preprint arXiv :1707.06742*, 2017.
- [28] R. Sutton and A. Barto. *Reinforcement learning : An introduction*. MIT press, 2018.
- [29] K. Van den Bos, R. Vermunt, and H. A. Wilke. The consistency rule and the voice effect : The influence of expectations on procedural fairness judgements and performance. *European Journal of Social Psychology*, 26(3) :411–428, 1996.
- [30] E. Wall, S. Ghorashi, and G. Ramos. Using expert patterns in assisted interactive machine learning : A study in machine teaching. In *IFIP Conf. on Human-Computer Interaction*, pages 578–599. Springer, 2019.
- [31] M. Yin, J. Wortman Vaughan, and H. Wallach. Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–12, 2019.
- [32] W. Younes. *Un système multi-agent pour la composition logicielle opportuniste en environnement ambiant et dynamique*. These de doctorat, Université Paul Sabatier - Toulouse III, June 2021.
- [33] R. Zhang, F. Torabi, G. Warnell, and P. Stone. Recent advances in leveraging human guidance for sequential decision-making tasks. *Autonomous Agents and Multi-Agent Systems*, 35(2) :1–39, 2021.