



Science Arts & Métiers (SAM)

is an open access repository that collects the work of Arts et Métiers Institute of Technology researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <https://sam.ensam.eu>
Handle ID: <http://hdl.handle.net/10985/22745>



This document is available under CC BY-NC-ND license

To cite this version :

Nicolas PERRY, Marc VAUTIER, Guillaume BUSATO, D. CHARLET, M. BACCOUCHE, G. ANTIPOV - IA ET EMPREINTE ENVIRONNEMENTALE : QUELLE CONSOMMATION D'ÉNERGIE POUR QUELLES ÉTAPES ? - 2022

Any correspondence concerning this service should be sent to the repository

Administrator : scienceouverte@ensam.eu



IA ET EMPREINTE ENVIRONNEMENTALE ; QUELLE CONSOMMATION D'ÉNERGIE POUR QUELLES ÉTAPES ?

Rapport de Projet Collaboratif du réseau EcoSD

Mars 2022⁽¹⁾

N° de contrat :
Étude réalisée pour le compte de l'ADEME par : *Non renseigné*⁽¹⁾

Coordination technique ADEME : *ERWAN Autret* – Direction\Service : *ANGERS DECD SPEM*



SYNTHÈSE D'ÉTUDE

En partenariat avec :



REMERCIEMENTS

Citer les membres du Comité de pilotage et/ou du comité de suivi ou de relecture et/ou du consortium de recherche.

Nicolas Perry (Arts et Métiers Bordeaux)
Marc Vautier (Orange)
Guillaume Busato (RTE)

CITATION DE CETTE SYNTHÈSE

Auteur(s) :

Orange : D.Charlet, M.Baccouche, G.Antipov, L.Charreire, V.Lemaire, P.Rust, L.Arga, M.Vautier,

RTE : G.Busato,

ENSAM : N.Perry, T.Durand, U.Paila, E.Abisset

Année de publication. Mars 2022

Titre du rapport : GREEN IA : Rapport de Projet Collaboratif du réseau EcoSD

En français :

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite selon le Code de la propriété intellectuelle (art. L 122-4) et constitue une contrefaçon réprimée par le Code pénal. Seules sont autorisées (art. 122-5) les copies ou reproductions strictement réservées à l'usage privé de copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations justifiées par le caractère critique, pédagogique ou d'information de l'œuvre à laquelle elles sont incorporées, sous réserve, toutefois, du respect des dispositions des articles L 122-10 à L 122-12 du même Code, relatives à la reproduction par reprographie.

En anglais :

Any representation or reproduction of the contents herein, in whole or in part, without the consent of the author(s) or their assignees or successors, is illicit under the French Intellectual Property Code (article L 122-4) and constitutes an infringement of copyright subject to penal sanctions. Authorised copying (article 122-5) is restricted to copies or reproductions for private use by the copier alone, excluding collective or group use, and to short citations and analyses integrated into works of a critical, pedagogical or informational nature, subject to compliance with the stipulations of articles L 122-10 – L 122-12 incl. of the Intellectual Property Code as regards reproduction by reprographic means.

Table des matières

1. Contexte du projet.....	4
2. Principaux résultats obtenus.....	5
2.1 IA : principes et mesure.....	5
2.2 Mesure des consommations d'énergie d'une IA.....	7
2.3 Dépendance aux algorithmes.....	9
2.4 Dépendance au contexte.....	11
3. Recommandations.....	11
4. Conclusions / Perspectives.....	12
Références bibliographiques.....	13
Sigles et acronymes.....	13
Annexes.....	14
Annexe 1.....	14
Annexe 2.....	14
Annexe 3.....	14
Annexe 4.....	14

1. Contexte du projet

Le développement d'approches d'Intelligence Artificielle (IA) dans de nombreux domaines et pour de nombreuses applications interroge, au-delà des enjeux de performances techniques et économiques, sur ses conséquences environnementales. Intégré dans la croissance accélérée du numérique, cette étude cherche à analyser les principaux éléments à retenir pour évaluer les impacts environnementaux de l'IA, en se concentrant sur commencer sur sa consommation énergétique.

NégaWatt (1) estime que la consommation électrique du numérique devrait augmenter de 15 TWh d'ici à 2030, soit +25 % par rapport à 2015 et porterait ainsi la part du numérique à environ 15 % de la consommation électrique du pays.

Le numérique et l'intelligence artificielle (IA) sont présentés comme des fondamentaux du monde à venir, en particulier en cette période de recours massif aux outils numériques.

Or de nombreux travaux (Shift Project [2][3], Flipo [4], Commissariat général au développement durable [5]) ont pu montrer que les différentes options de développement du numérique sont loin d'être parées des mêmes vertus de soutenabilité. Flipo analyse même que « les effets écologiques sont plus que compensés par l'expansion rapide du secteur et la digitalisation des modes de vie qui va avec. D'où notamment des émissions de GES qui doublent tous les quinze ans, sans que cette hausse soit compensée par un effet positif sur les autres secteurs. »

Le cas particulier de l'IA et de ses impacts environnementaux en regard de ses performances, est un sujet qui émerge dans la communauté IA mais qui reste encore peu abordé et qui n'a pas de consensus sur une méthode et un périmètre d'évaluation.

Dans un souci d'efficacité et consciente de l'ensemble des impacts à étudier, notre proposition commencera par s'intéresser aux consommations énergétiques. Les aspects matériaux, eux, seront ici peu abordés bien qu'ils constituent une suite naturelle à ces travaux.

L'objectif de ce travail est de défricher ce sujet en ayant pour objectifs principaux de :

1. faire œuvre de pédagogie sur l'IA
2. présenter les analyses des différentes étapes d'IA (constructions/apprentissage et usage/inférence) sous l'angle consommation d'énergie, précision de calcul mettre en avant les principales conclusions sur les outils de mesure,

Cette analyse a été menée en considérant deux contextes d'études différents et complémentaires apportés par les partenaires :

1. TIC : le cas d'étude proposé par Orange est l'outil interne Orange Cobalt ASR (Automatic Speech Recognition) de reconnaissance vocale.
2. Gestion de l'énergie : RTE utilise des algorithmes d'IA pour équilibrer la consommation et production d'énergies sur le réseau électrique, en se basant sur des prédictions et des demandes temps réelles en adéquation avec les capacités de productions des différents types d'énergies. Cet équilibrage se fait à des mailles de réseaux géographiques différentes (régional, national voir international). Ce cas n'a pas pu être traité dans le cadre de ce projet et pourra faire l'objet de développements futurs.

La nature des contextes, des données manipulées, des attendus en temps de réponse, des niveaux de précision, et des tolérances aux erreurs sont de natures très diverses.

La comparaison des contextes et des approches alimente le travail, et permet d'extrapoler vers des enjeux spécifiques et des problématiques à lever pour aborder lucidement l'analyse d'une solution d'IA mise en place et accompagner l'aide au choix d'une IA énergétiquement sobre .

Le détail des résultats sont disponibles dans les Annexes

- 1 : Introduction to Green IA - ORANGE 2020,
- 2 : IA et Sobriété - ENSAM 2021
- 3 : Green AI – ENSAM 2021
- 4 : Measuring the energy costs of an AI-based software – ORANGE 2022

2. Principaux résultats obtenus

L'objectif de cette étude est d'analyser la consommation énergétique des algorithmes et processus d'Intelligence Artificielle et d'analyser les différents aspects qui peuvent les impacter.

Les principales leçons sont les suivantes:

- Les algorithmes d'IA se construisent et se mettent en œuvre en plusieurs étapes (capture données/traitement/stockage, apprentissage, inférences (usage)). Ces différentes étapes peuvent en fonction des IA étudiés avoir des poids énergétiques différents. Ainsi en fonction du format des données recueillies et à traiter, différentes étapes de traitement/nettoyage peuvent être nécessaires (données provenant de capteurs par exemple) avant de lancer les procédures d'apprentissage, dont certaines peuvent ne durer que quelques heures et tourner sur des machines simples, alors que d'autres nécessitent des ressources importantes et des durées d'apprentissage eux aussi très long. Enfin pour la phase inférence ou usage, une application utilisée quelques fois par jour mais par des milliers/millions de personnes (enceinte vocale) nécessitera au final plus de ressource qu'une application avec un usage globalement beaucoup plus faible. L'identification de la phase la plus impactante d'un point de vue consommation de ressources doit donc être menée avec attention.

- Une précision élevée dans la recherche des résultats ne nécessite pas obligatoirement un nombre d'apprentissage important. Le ratio précision / consommation d'énergie (et donc émissions environnementales) doit être considéré.

- Le coût environnemental de l'IA est relatif au coût évité du système physique sur lequel opère et traite les informations. Ainsi il faut pouvoir mesurer l'effet bénéfique de l'algorithme développé à l'aune de l'effet sur le système.

- Il y a une forte dépendance de la consommation d'énergie aux algorithmes d'apprentissage et de traitement.

- Un travail important est à mener sur la fiabilisation de la mesure de la consommation d'énergie de différents process tournant sur des machines (sondes logicielles...).

2.1 IA : principes et mesure

Définition

Pour avoir une présentation complète sur l'Intelligence Artificielle, il est fortement recommandé de découvrir ces technologies sur le site cours en ligne gratuit "Elements of AI" (<https://www.elementsofai.fr/>).

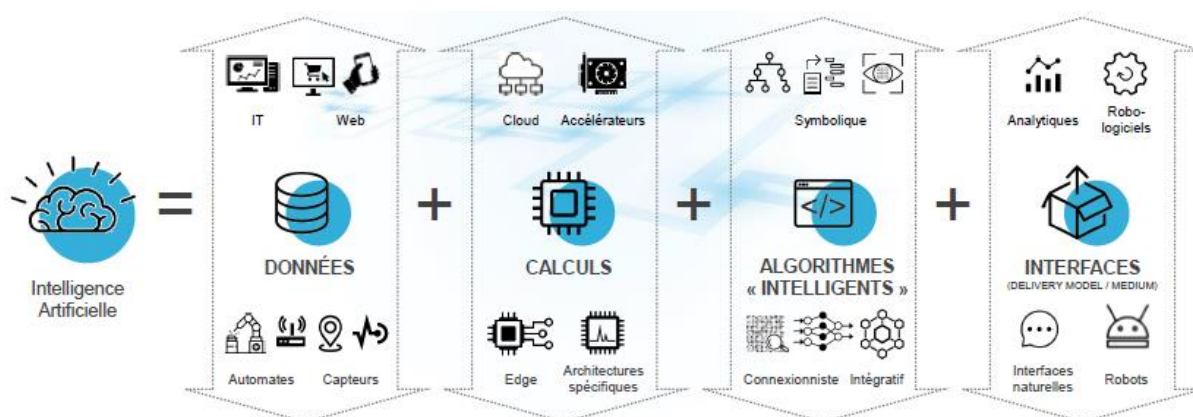


Figure 1 : illustration des différentes facettes d'une IA (source : xxx)

Une IA peut être vue comme un système robotique, soit virtuel soit physique, qui analyse des données/signaux et produit un résultat, qu'il soit cognitif (savoir, connaissance, décision...) ou moteur (manipulation, navigation, interaction...) comme illustré sur la Figure 1.

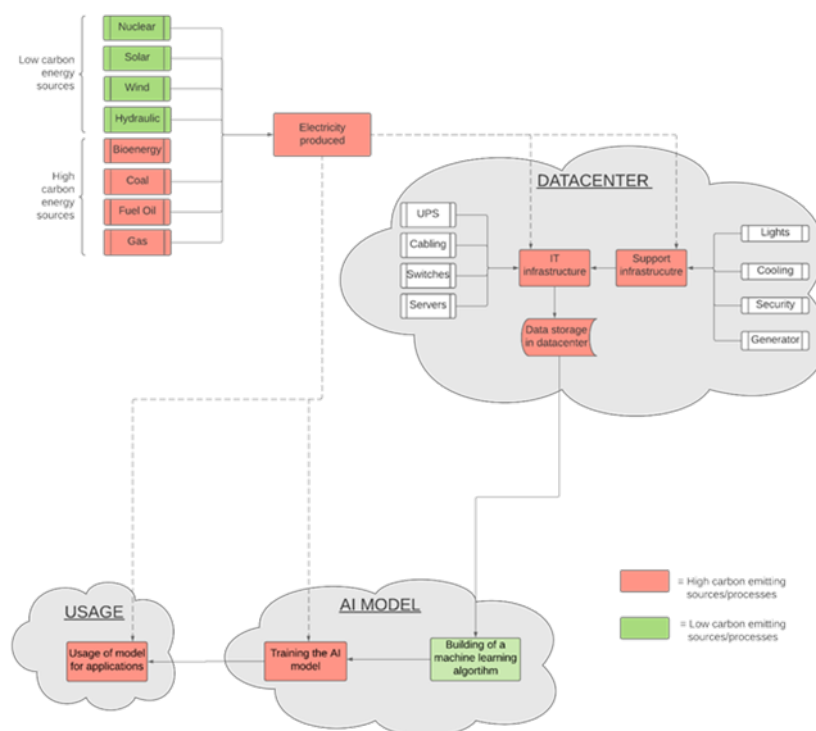


Figure 2 : Exemple d'architecture de stockage d'une IA

De plus il faut retenir qu'il y a un processus pour aboutir à une solution d'IA compatible au contexte d'usage en lien avec la précision attendue et les nombre de données (data set) disponible pour construire l'algorithme.

Il y a donc :

1. une étape de **construction** du jeu de données d'entraînement (**Data Set**),
2. une étape d'**apprentissage** (entraînement ou **training**), qui peut être supervisé ou non supervisé.

Une étape complémentaire de **renforcement** peut être appliqué pour atteindre un niveau de précision supérieur en jouant sur le nombre d'**hyperparamètres**. Sachant qu'en apprentissage automatique, un hyperparamètre est un paramètre dont la valeur est utilisée pour contrôler le processus d'apprentissage.

3. une étape de **déploiement (usage/inférence)**.

L'ensemble de ses étapes nécessite du **stockage (Data Center)** et des échanges de données, comme illustré sur la Figure 2.

La construction des données de départ et l'entraînement, en fonction des algorithmes utilisés vont conditionner la précision de l'IA ainsi construite.

Chaque étape consomme des ressources (énergie/matière) pour le calcul ou le stockage des données, conduisant ainsi à des impacts environnementaux. Dans la présente étude, le focus est fait sur la consommation d'énergie sur les phases d'apprentissage et inférence. La consommation de matière n'a pas été considérée.

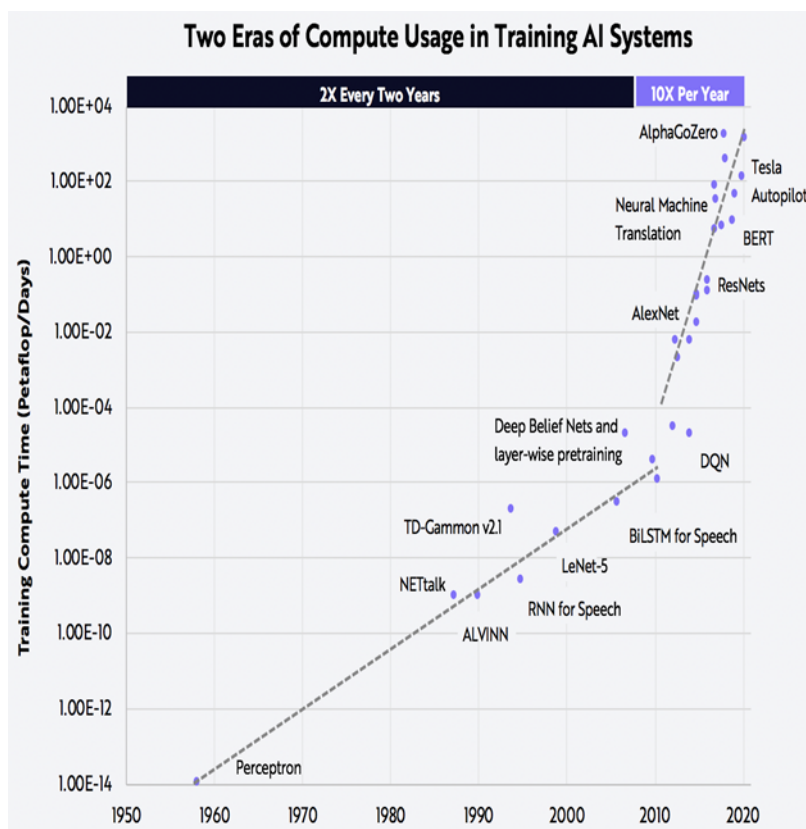


Figure 3 : Evolution of ML systems' computational needs: doubling every two years from 1960 to 2010. 10x per year since 2010 (Wang 2020)

Les unités centrales de traitement (CPU) et les unités de traitement graphique (GPU) sont des moteurs de calcul essentiels au fonctionnement des machines qui supportent les technologies d'IA. L'évaluation de la consommation de ressources énergétiques d'une IA revient donc dans une première approximation à celle des CPU/GPU, ce qui demande la mise en œuvre d'outils spécifiques comme les sondes logicielles pour la mesurer.

Chaque contexte est spécifique et peut conduire à des distributions d'impacts propres à la combinaison Data Set / algorithme d'IA / précision ciblée / inférence.

Plusieurs approches d'IA existent sous le terme générique de Machine Learning : les réseaux neuronaux, les régressions logistiques, les K-voisins, l'arbre de décision ... (voir Figure 3). Ces différentes approches sont présentées dans l'annexe 2 §6.a.

2.2 Mesure des consommations d'énergie d'une IA

Les travaux récents de chercheurs et d'acteurs de l'IA (voir Annexe 1 # 3.2) font ressortir les indicateurs suivants à évaluer :

- la quantité de données nécessaire pour entraîner le modèle de machine learning,
- le nombre de séquences nécessaires pour atteindre la précision souhaitée par apprentissage profond
- la quantité d'énergie nécessaire à traiter une inférence,

Une connaissance précise de la quantité d'énergie des différents process d'une IA a nécessité de passer par une étape de mesure de cette consommation d'énergie au niveau des équipements utilisés. Dans le cas présent la mesure de cette énergie sera celle des CPU ou GPU utilisées, comme décrits dans l'annexe 1#3.3.

Il faut noter que la mesure de l'énergie est dépendante de la machine et du matériel employés.

Les détails pour accéder aux données relatives au CPU (processeur de calcul) et GPU (processeur Graphique) sont décrits dans l'annexe 1 # 3.3.

Ces consommations d'énergies sont traduites en émissions de CO₂e et empreinte carbone.

4 sondes open sources ont été identifiées et testées, elles ont permis de faire ces mesures de consommation d'énergie :

- Power API (co développé par Orange et INRIA),
- PyJoules,
- Scaphandre,
- CodeCarbon (voir Annexe 4 #3).

Il existe aussi un site web de calculateur d'émission pour l'apprentissage d'une IA (ML CO₂ Impact :

<https://mlco2.github.io/impact/#compute>).

Le tableau 1 donne les principales caractéristiques des sondes logicielles testées :

Tool	Power API	pyJoules	Scaphandre	CodeCarbon
Developers	Spiral Research Group (INRIA) + contribution from Orange Innovation (IT-S)	Part of the Power API initiative	Hubblo.org initiative	MILA, BCG Gamma and Haverford College
Scope	CPU only	CPU + GPU	CPU only	CPU + GPU
Measures granularity	Per process	Global	Per process	Global
Data Sources	RAPL + ML on hardware/OS counters	RAPL + NML	RAPL	RAPL + NML
OS Support	Linux only	Linux only	Linux only	Linux, Windows, MacOS

Tableau 1 : Résumé des principales caractéristiques de sondes open-sources

Les conclusions de ces comparaisons d'outils ont été obtenues sur la base de jeux de données similaires (reconnaissance vocale).

Ces sondes sont d'une manière générale encore au stade de développement (plus ou moins avancés) et les résultats fournis doivent être pris avec précaution. Scaphandre n'a pas été menée au bout des essais par difficulté de configuration.

Les principales conclusions sont :

- Pour la phase d'inférence :
 - les résultats de Power API et pyJoules sont cohérents et très similaires
 - le logiciel CodeCarbon sous-estime d'un facteur 1000 les mesures, en partant du principe que les résultats Power API font référence.
- Pour la phase d'apprentissage :
 - Les logiciels pyJoules et CodeCarbon donnent des évaluations comparables
 - Le ratio d'usage des ressources CPU – GPU n'étant pas 100 % sur l'intégralité des temps d'apprentissage, il est difficile d'estimer a priori la part du temps respectif de traitement CPU et GPU.

Commentaires : D'une manière générale un travail important est à mener sur la fiabilisation de la mesure de la consommation d'énergie de différents process tournant sur des machines.

L'amélioration des précisions des résultats de l'inférence n'est pas linéairement dépendante du nombre d'itérations, alors que la consommation en ressource l'est : il y a donc un optimum à trouver entre précision et consommation énergétique. Ainsi, la figure suivante montre qu'après 50 blocs d'apprentissages la précision des résultats n'évolue plus.

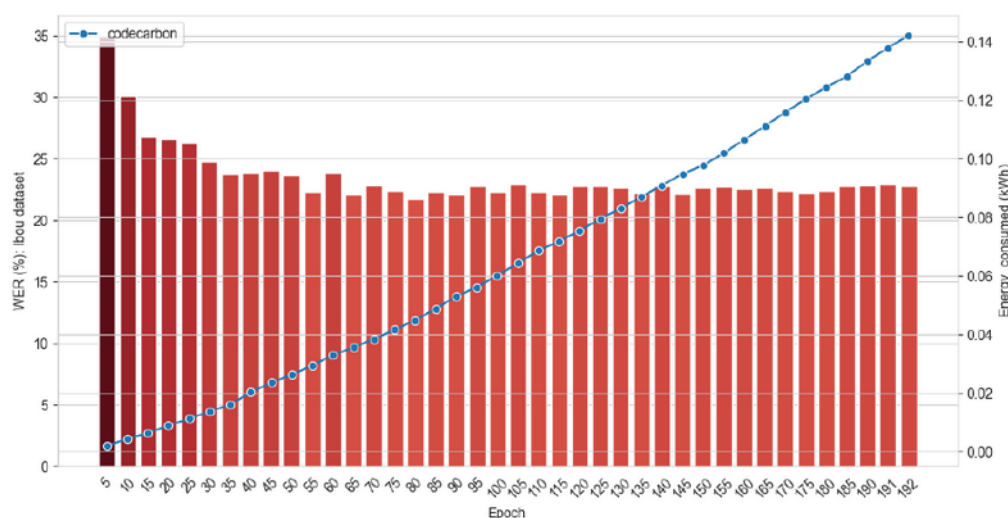


Figure 4 : Consommation d'énergie cumulée (kwh mesuré par CodeCarbon) fonction de la précision (en % WER sur la base de donnée Ibou) évaluées à différentes itérations de l'apprentissage du code Cobalt-ASR (application à la reconnaissance vocale)

2.3 Dépendance aux algorithmes

L'objectif de cette partie est de comparer pour différents algorithmes et jeux de données les consommations d'énergie obtenues ainsi que la précision des résultats. Pour plus de détails voir l'Annexe 2§6.C.

Les modèles suivants ont été considérés (pour plus d'information voir Annexe 2§6.a :

- Réseau neuronal
- Régressions logistique
- K-voisins
- Arbre de décision

Les modèles de données utilisés :

Wine

- 260kB 4900 entrées 11 caractéristiques par entrée 7 classes de sorties

Ce jeu de données contient des informations sur 4900 vins blancs. Ces vins sont associés à une note de qualité attribuée par un testeur.

Mushrooms

- 1150kB 8120 entrées 21 caractéristiques par entrée 2 classes de sorties

Le dataset Mushrooms rassemble diverses caractéristiques de taille et de forme de plus de 8000 champignons. Ces champignons sont séparés en deux catégories en fonction de leur comestibilité.

Mnist

- 107MB 60000 entrées 784 caractéristiques par entrée 10 classes de sorties

Mnist est un des datasets les plus connus dans le monde du machine learning. Il est très largement utilisé afin de tester le fonctionnement d'un algorithme à petite échelle. Il contient des images en nuances de gris de 60 000 chiffres écrits à la main et de résolution 28x28 pixels. Chaque image est stockée sous la forme d'une liste d'entiers entre 0 et 255 représentant la luminosité de chaque pixel de l'image. A chaque image est associé le chiffre qu'elle représente.

Mnist fashion

- 130MB 60000 entrées 784 caractéristiques par entrée 10 classes de sorties

Mnist fashion est très similaire à Mnist. Ici ce ne sont pas des chiffres mais des pièces vestimentaires qui ont été photographiées et réparties dans 10 catégories. Ce jeu de données est souvent utilisé comme alternative à Mnist.



Figure 5- Résultats de mesures de consommation d'énergie & précision pour différentes méthodes d'IA et jeux de données

Les graphes ci-dessus présentent une synthèse des mesures énergétiques & de précision obtenues pour les différents jeux de données traités avec les 4 types d'algorithmes d'IA. Il en ressort que dans toutes les situations étudiées le classement des algorithmes du point de vue de la consommation électrique reste identique. Le modèle le plus économe est l'arbre de décision qui consomme entre 5 et 40 fois moins que le modèle des k-voisins qui se place en deuxième

position. L'algorithme le plus gourmand de loin est le réseau neuronal qui consomme en moyenne 4000 fois plus que l'arbre de décision. En termes de précision cependant le réseau neuronal passe en tête. Il est en effet le plus précis dans 3 jeux de données sur 4 et est en moyenne 6% plus précis que le deuxième meilleur algorithme pour chaque jeu de données.

Dans ce classement l'arbre de décision se place en dernière position dans tous les tests et est en moyenne 25% moins précis que le réseau neuronal, avec dans le cas de Mnist et Mnist fashion une précision de 60% quand tous les autres algorithmes restent au-dessus de 95%

Concernant les infrastructures indissociables des solutions d'IA, les impacts sont similaires à ceux liés au numérique et aux TIC.

On retrouve ainsi une forte dépendance au stockage de données et aux data center avec des projections de développement de leur nombre du fait d'un usage plus large et plus généralisé des technologies du numérique. Les technologies d'intelligence artificielle y contribuent elles aussi.

2.4 Dépendance au contexte

Pour juger de l'intérêt et du coût environnemental d'une application d'IA, il est nécessaire de connaître le cas d'usage et son incidence environnementale. Une approche coût/ bénéfice apporte alors une vision sur le consentement à payer entre autres lié à la précision de la réponse.

Il faut donc évaluer conjointement le gain environnemental du cas étudié (associé par exemple au temps de réponse ou à la précision de traitement, ou à d'autres gains en fonction du service rendu) avec le coût environnemental de la solution.

3. Recommandations

Les recommandations suivantes représentent des pistes d'amélioration des performances environnementales d'une IA.

Données

- Evaluer les impacts des différents types d'architecture liés à la collecte, au transfert et au traitement des données. En effet, les différentes architectures de calcul possibles (Cloud computing / Edge computing / Local computing) impactent les besoins en échanges de données et en infrastructures. Ainsi tout calcul local réduit les temps de réponse (latences) mais aussi les consommations indirectes en ressources ramenées au service d'IA. Toutefois un calcul local peut nécessiter une multiplication d'outils au plus proche du phénomène étudié.
- Objectiver les conséquences des choix d'emplacement des Data centers. Ainsi, les Data Center et les centres de calculs peuvent être optimisés en termes de positionnement géographique vis à vis du mix énergétique du pays ou de la région d'implantation. Ce choix d'implantation centré énergie et émissions de CO2 est à mettre en balance des éventuels besoins en infrastructures et flux générés sur le réseau. Il faut évaluer le possible risque de transfert d'impact lié aux distances plus importantes et donc aux impacts cachés du fait des énergies et ressources nécessaires pour assurer le complément d'infrastructure.

Algorithmes

- Réduire le nombre d'Hyperparamètres et chercher le niveau de complexité réduit qui reproduit suffisamment le problème à traiter.
- Limiter les itérations et la taille du jeu de données à traiter, car le nombre d'itérations d'apprentissage conditionne directement la consommation d'énergie.
- Bien définir le niveau de précision souhaité en fonction du contexte du cas d'application et des algorithmes utilisés. La précision étant asymptotique, le nombre d'itérations peut être optimisé lors des phases d'apprentissage.

Méthode

- Evaluer conjointement le bénéfice environnemental de l'usage d'une IA (pour bien statuer sur les exigences en précision par exemple) et le coût environnemental de cette construction d'IA et statuer par exemple sur le nombre d'itération pour l'apprentissage ou le nombre d'hyperparamètres pour représenter la complexité des systèmes.

4. Conclusions / Perspectives

Ce projet a eu pour modeste ambition d'initier des investigations sur le cout énergétique de la mise en œuvre d'une intelligence artificielle et de présenter la complexité de la mesure de cette énergie par des sondes logicielles. L'ensemble de ces travaux devront être repris pour par exemple caractériser en détails le cout énergétique des différentes étapes représentatives d'une IA (collecte & traitement des données, apprentissage, inférence), avec différents types d'algorithmes et différentes précisions.

Ce travail trouve aussi un écho dans la loi REEN relative à la réduction de l'empreinte environnementale du numérique en France (8)(9) qui se décline selon 5 axes :

- Sensibiliser et faire prendre conscience de l'impact environnemental du numérique,
- Limiter le renouvellement des appareils numériques,
- Favoriser des usages numériques écologiquement vertueux,
- Promouvoir des datacenters et des réseaux moins énergivores,
- Promouvoir une stratégie numérique responsable dans les territoires.

Cette réglementation a tout de même des limites sur les aspects hardware et matériel, avec peu d'obligation pour les fabricants à travailler sur l'allongement de la durée de vie, d'un part, et d'autre part, peu de propositions apportant des avantages au matériel reconditionné.

Références bibliographiques

1. Negawatt (2020) Numérique : peut-on passer de la gabegie à la sobriété ?, Université Negawatt 2020 - webinaire replay de session
2. Shift Project (2018) Lean ICT. Pour une sobriété numérique. Rapport du Shift Project.
3. Shift Project (2021) Impact environnemental du numérique : tendances à 5 ans et gouvernance de la 5g. Rapport du Shift Project
4. Flipo, F. (2021). L'impératif de la sobriété numérique. *Cahiers Droit, Sciences & Technologies*, (13), 29-47.
5. GDS EconInfo (2019), Data centres, Conséquences du numérique dans le site <https://ecoinfo.cnrs.fr/>
6. Commissariat général au développement durable (2019) Numérique et consommation énergétique, sur site de notre-environnement.gouv.fr
7. J.Wang (2020), The Cost of AI Training is Improving at 50x the Speed of Moore's Law: Why It's Still Early Days for AI, Ark Invest, <https://ark-invest.com/analyst-research/ai-training/>
8. Assemblée Nationale (2021) LOI n° 2021-1485 du 15 novembre 2021 : réduire l'empreinte environnementale du numérique en France, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044327272>
9. <https://www.vie-publique.fr/loi/278056-loi-15-novembre2021-reen-reduire-empreinte-environnementale-du-numerique>

Sigles et acronymes

IA	Intelligence Artificielle
ML	Machine Learning
CPU	unités centrales de traitement : processeur
GPU	unités de traitement graphique : processeur graphique

Annexes

Annexe 1

Introduction to Green AI : A reading grid of Machine Learning's environmental aspects

Auteurs : Delphine CHARLET Moez BACCOUCHE

Keywords : Artificial Intelligence, green, energy consumption, training

Annexe 2

Etude préparatoire sur les impacts environnementaux des systèmes d'Intelligence Artificielle (IA)

Auteur : T. Durand janv 2021

Stage Master

Annexe 3

Green AI

Auteur : U.PAILA Sept 2021

Stage Master

Annexe 4

Measuring the energy costs of an AI-based software : Case study on the training and inference phases of Cobalt-ASR

Auteurs : Grigory ANTIPOV Moez BACCOUCHE Delphine CHARLET Laurent CHARREIRE
Vincent LEMAIRE Pierre RUST Ludovic ARGA

Keywords : Green AI, responsible AI, energy consumption measurement, carbon emissions, deep learning

L'ADEME EN BREF

L'Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME) participe à la mise en œuvre des politiques publiques dans les domaines de l'environnement, de l'énergie et du développement durable. Elle met ses capacités d'expertise et de conseil à disposition des entreprises, des collectivités locales, des pouvoirs publics et du grand public, afin de leur permettre de progresser dans leur démarche environnementale. L'Agence aide en outre au financement de projets, de la recherche à la mise en œuvre et ce, dans les domaines suivants : la gestion des déchets, la préservation des sols, l'efficacité énergétique et les énergies renouvelables, la qualité de l'air et la lutte contre le bruit.

L'ADEME est un établissement public sous la tutelle conjointe du ministère de l'Ecologie, du Développement durable et de l'Energie, et du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.



ADEME
20, avenue du Grésillé
BP 90406 | 49004 Angers Cedex 01

   .fr
ADEME
20, avenue du Grésillé
BP 90406 | 49004 Angers Cedex 01

ABOUT ADEME

The French Environment and Energy Management Agency (ADEME) is active in the implementation of public policy in the areas of the environment, energy and sustainable development. The Agency provides expertise and advisory services to businesses, local authorities and communities, government bodies and the public at large, to enable them to establish and consolidate their environmental action. As part of this work ADEME helps finance projects, from research to implementation, in the areas of waste management, soil conservation, energy efficiency and renewable energy, air quality and noise abatement.

ADEME is a public agency under the joint authority of the Ministry for Ecology, Sustainable Development and Energy, and the Ministry for Education, Higher Education and Research.



ADEME
20, avenue du Grésillé
BP 90406 | 49004 Angers Cedex 01

www.ademe.fr