MINES Saint-Étienne

Institut Mines-Télécom

Apprentissage Fédéré

Vers une IA plus respectueuse de la vie privée et moins consommatrice d'énergie

INSPIRING INNOVATION SINCE 1816

Guillaume MULLER

Qui suis-je ?

- **Formation**: Dr en informatique (Multi-Agent)
- **Enseignement** (@EMSE): Informatique, AI & Cyber-Security
- **Recherche** (@Fayol): Membre du dept. ISI

Mon parcours

- **P2P**
- **Assistant Virtuel** (2007!)
- **Jeu Sérieux**
- **Maison Solaire**
- **Moteur de Recherche**
- **Aviation**
- **Industrie 4.0**

Projets actuels

- **Apprentissage Fédéré, TinyML**

Numérique Responsable ?

3 piliers du Développement Durable

- **Économique**
  - Impacts sur l'emploi
  - ...
- **Social**
  - Questions **éthiques** / morales
  - ...
- **Écologique**
  - toutes **énergies** (pas que CO_2)
  - toutes les ressources (eau, matériaux, ...)
  - tout le cycle de vie : création, **usage**, destruction

Numérique Responsable ?

3 piliers du Développement Durable

- **Économique**
  - Impacts sur l'emploi
  - ...
- **Social**
  - Questions **éthiques** / morales ← **privacy**
  - ...
- **Écologique**
  - toutes **énergies** (pas que CO_2) ← **réduction**
  - toutes les ressources (eau, matériaux, ...)
  - tout le cycle de vie : création, **usage**, destruction

Qu'est-ce que l'Intelligence ?

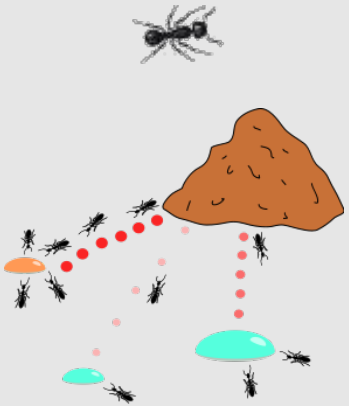
2 « vues »

- **Vue individuelle**
  - "Résoudre problème difficile/complexe"

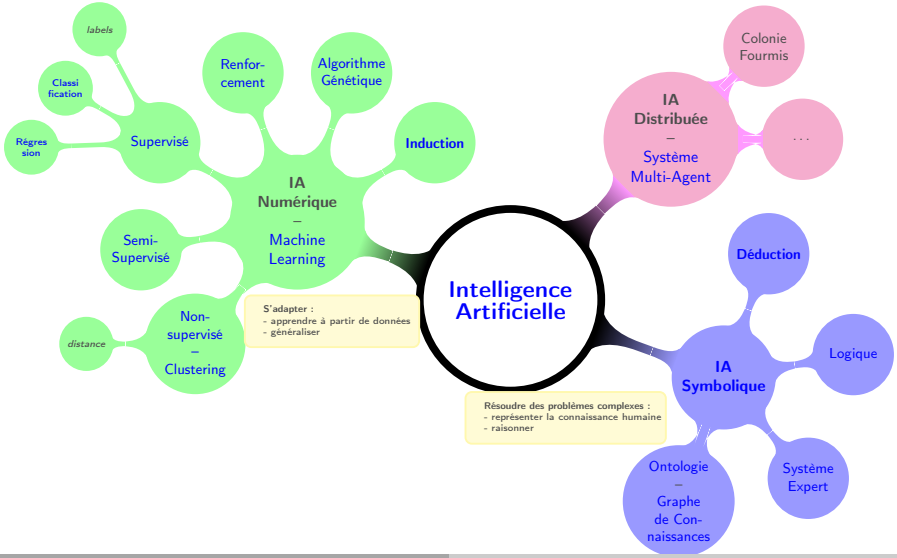
Qu'est-ce que l'Intelligence ?

2 « vues »

- **Vue individuelle**
  - "Résoudre problème difficile/complexe"
- **Vue sociétale**
  - Darwin : théorie évolution
  - Dawkins : "gène égoïste"
  - Survie de l'espèce par l'**Adaptation**
  - L'adaptation individuelle par l'**Apprentissage**

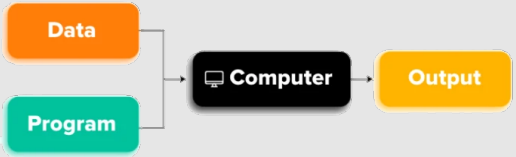


Qu'est-ce que l'Intelligence Artificielle (IA) ?



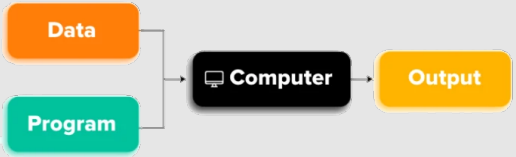
Apprentissage Machine ("Machine Learning" – ML)

Programmation Traditionnelle

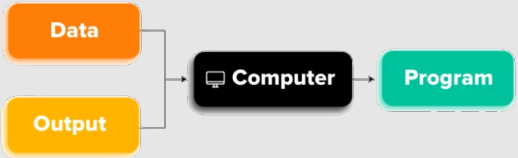


Apprentissage Machine ("Machine Learning" – ML)

Programmation Traditionnelle



Intelligence Artificielle



Exemple d'Apprentissage Machine

3 types d'Iris

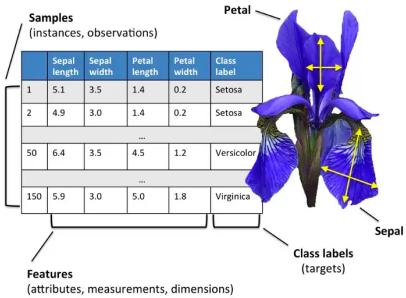
Iris Versicolor | Iris Setosa | Iris Virginica

Exemple d'Apprentissage Machine

3 types d'Iris

Iris Versicolor | Iris Setosa | Iris Virginica

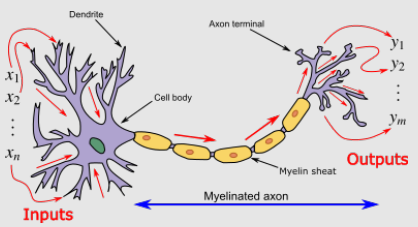
Collecte de données



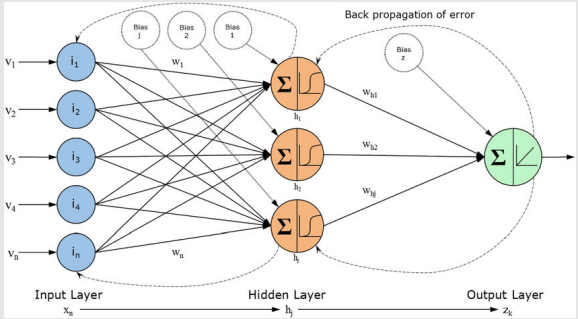
	Sepal length	Sepal width	Petal length	Petal width	Class label
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
...					
50	6.4	3.5	4.5	1.2	Versicolor
...					
150	5.9	3.0	5.0	1.8	Virginica

Exemple d'algorithme de ML: Réseau de neurones

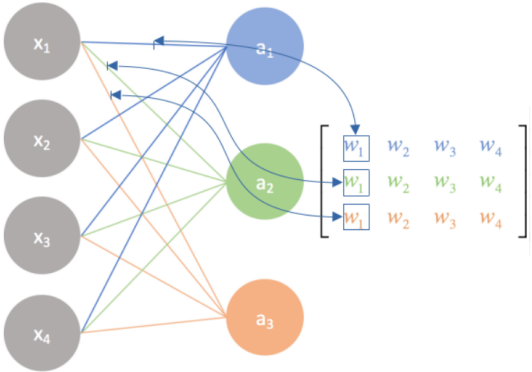
Biological Neuron



Artificial Neuron

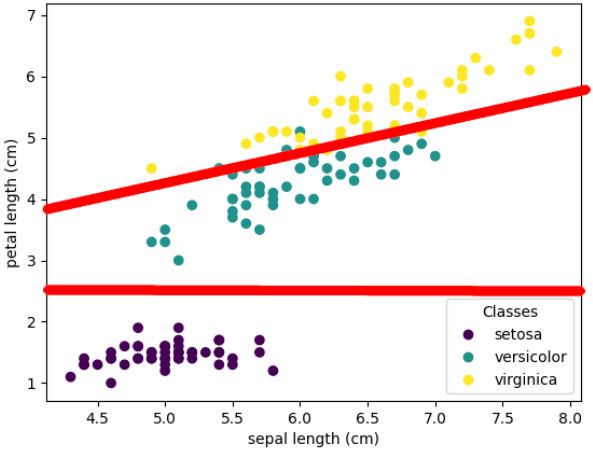


Réseau de neurones ∼ Matrice/Tenseur

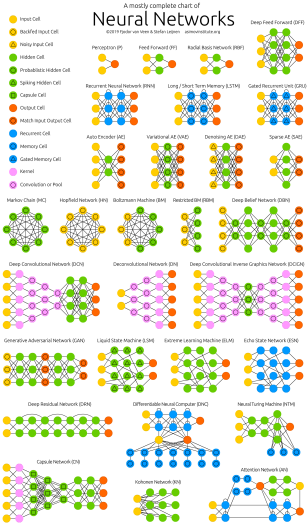


From: jeremyjordan.me

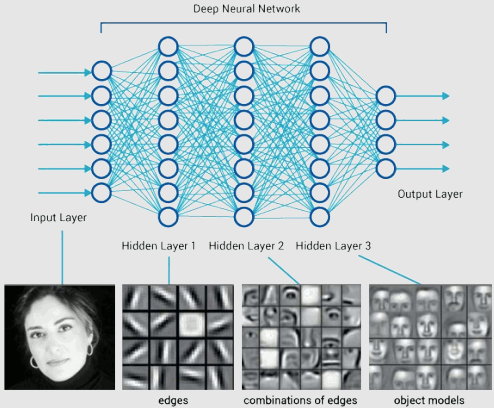
Solution


petal length (cm)
sepal length (cm)
Classes
setosa
versicolor
virginica

"Zoologie" de réseaux de neurones

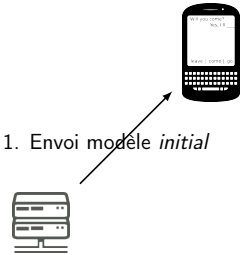


2ème Example : images



3ème Example : textes

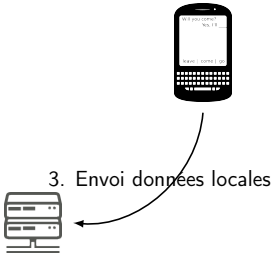
Étapes de l'apprentissage machine "Traditionnel"



Étapes de l'apprentissage machine "Traditionnel"

2. Utilisation modèle
   Collecte données

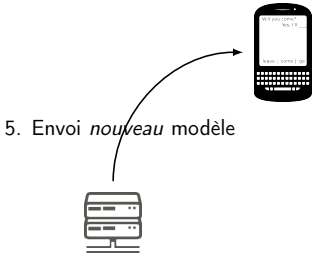
Étapes de l'apprentissage machine "Traditionnel"



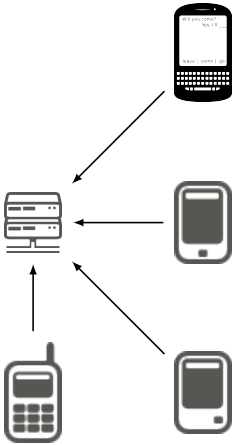
Étapes de l'apprentissage machine "Traditionnel"

4. Entraînement du modèle

Étapes de l'apprentissage machine "Traditionnel"



Étapes de l'apprentissage machine "Traditionnel"



Limites de l'apprentissage machine "Traditionnel"

Vie privée

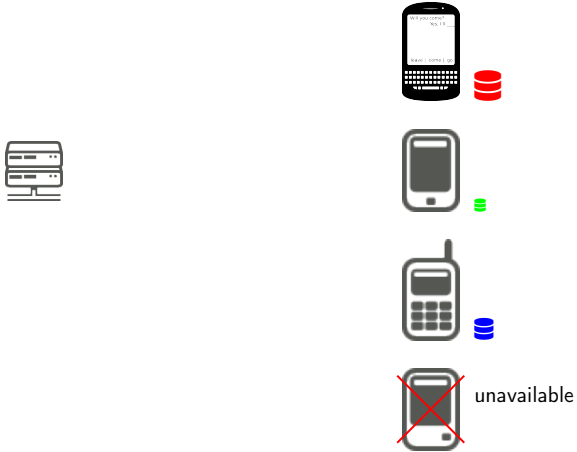
- Tapé au clavier :
  - Mots de passe
  - Messages privés
  - Sites web visités
  - ...

Efficience

- Données brutes = lourd ⚖

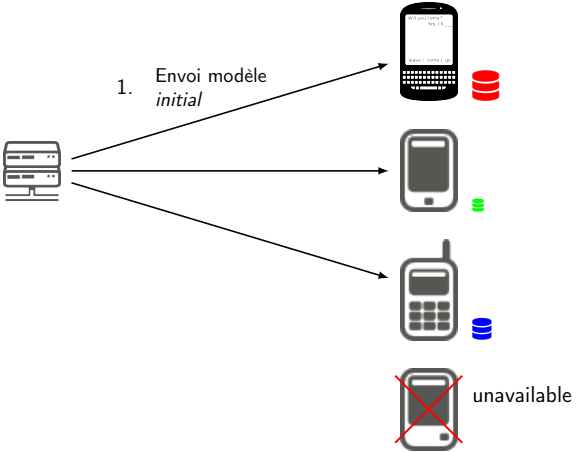
- texte	$\sim$ o/ko
- images	$\sim$ ko/Mo
- son	$\sim$ ko/Mo
- vidéos	$\sim$ Mo/Go

Apprentissage Fédéré – Principe



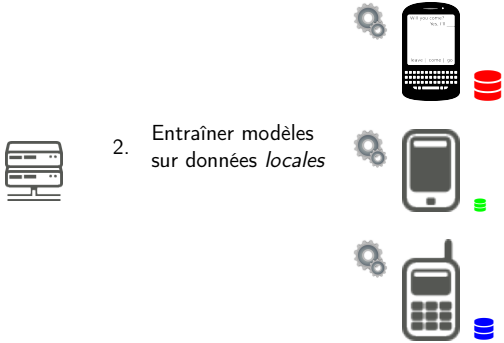
Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Principe



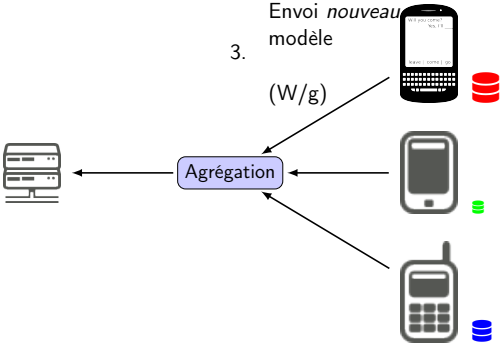
Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Principe



Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Principe



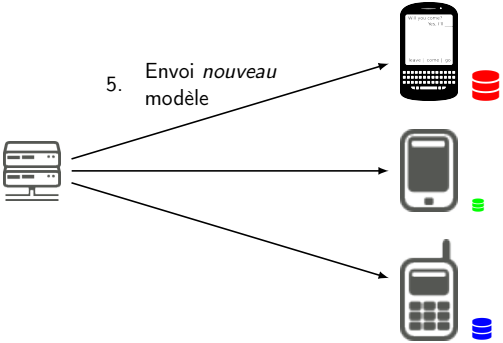
Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Principe

4. Nouveau modèle *global*

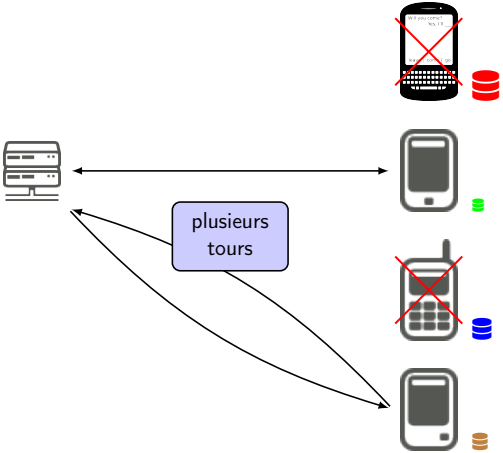
Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Principe



Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Principe



Source: H.B. McMahan *et al* (2017)

Apprentissage Fédéré – Avantages et Inconvénients

Avantages

- **Vie Privée**
  - nouveaux modèles = non "lisibles"*
- **Efficience**
  - **Processeur**: calculs distribués**
  - **Réseau**: `taille(données brutes)` $\ggg$ `taille(nouveaux modèles)`***

Inconvénients

- **Pratique** : ça marche!****
- **Théorie** : Encore *beaucoup* de recherche sur *beaucoup* d'aspects !

Problème de vie privée

Inversion de Modèle (M. Fredrikson *et al*, 2015)

Problème de vie privée

Inversion de Modèle (M. Fredrikson *et al*, 2015)

Problème de vie privée

Inversion de Modèle (M. Fredrikson *et al*, 2015)

Inversion de Modèle "protégé" (B. Hitaj *et al*, 2017)

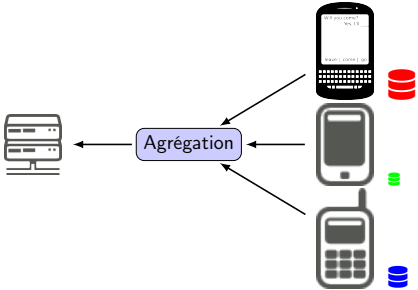
Original

$\theta_u = 1$
$\theta_d = 1$

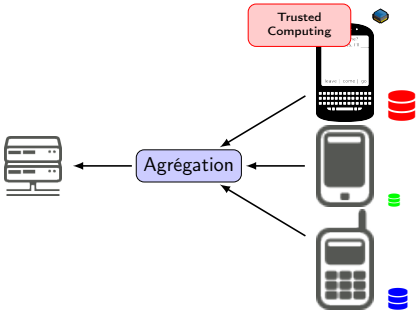
$\theta_u = 0.1$
$\theta_d = 1$

$\theta_u = 0.1$
$\theta_d = 0.1$

Protéger la Vie Privée en Fédéré (source: C. Wierzynski)



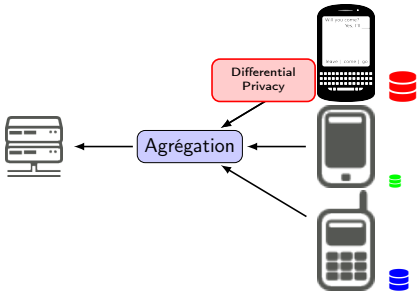
Protéger la Vie Privée en Fédéré (source: C. Wierzynski)



Trusted Execution Environment (TEE)

- Prevents external processes from reading RAM
  - Raw Data, Model, Gradients...
- Hardware protection (Intel, AMD, ARM, RISC-V...)

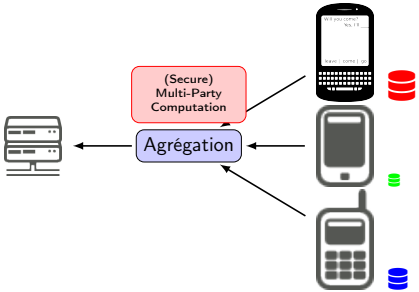
Protéger la Vie Privée en Fédéré (source: C. Wierzynski)



Differential Privacy (DP), (C. Dwork *et al*, 2014)

- Add noise to dataset so that stats remain same ⇒ ML possible
  - Limited Bugdet
- Limits "Model Inversion" attacks

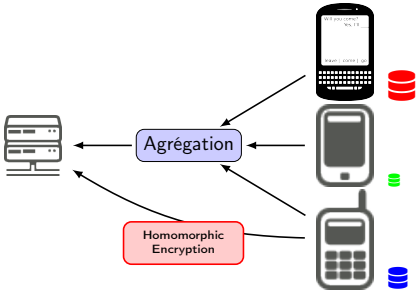
Protéger la Vie Privée en Fédéré (source: C. Wierzynski)



(Secure) Multi-Party Computation ([S]MPC)

- Jointly compute a function while keeping inputs private
  - $(\mathcal{A} + x) + (\mathcal{B} + y) + (\mathcal{C} + z) = (\mathcal{S} - s), s = x + y + z$
- Protects from network eavesdroppers (MitM)

Protéger la Vie Privée en Fédéré (source: C. Wierzynski)

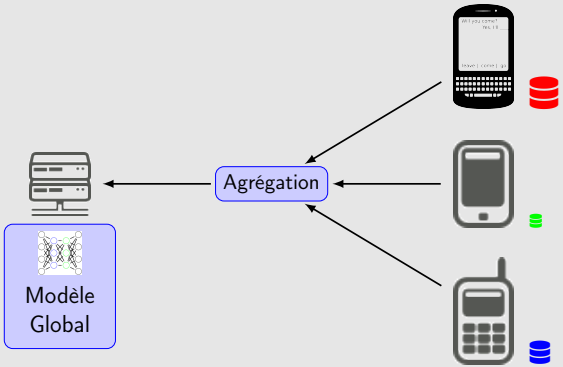


Homomorphic Encryption (FEH) (C. Gentry, 2009)

- Compute directly on cyphered data
- Similar technology as SMPC
- 🏆, but order of magnitude less efficient

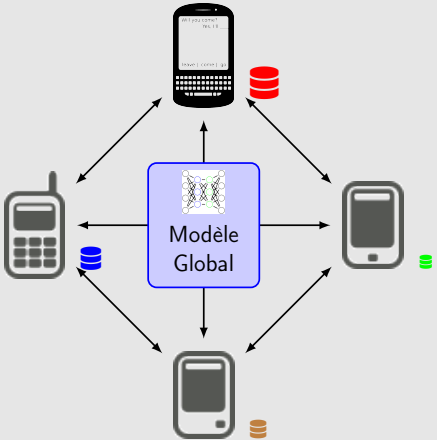
La recherche en Apprentissage Fédéré – Architecture

Centralized



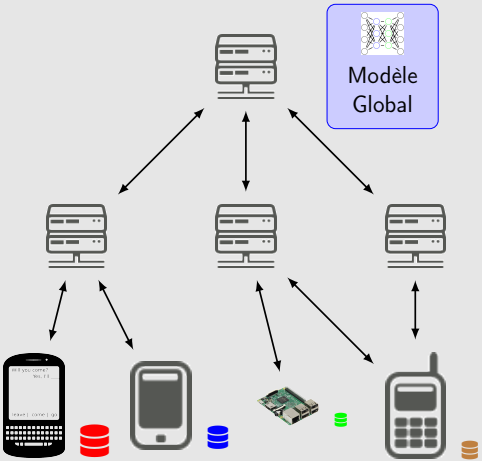
La recherche en Apprentissage Fédéré – Architecture

Decentralized

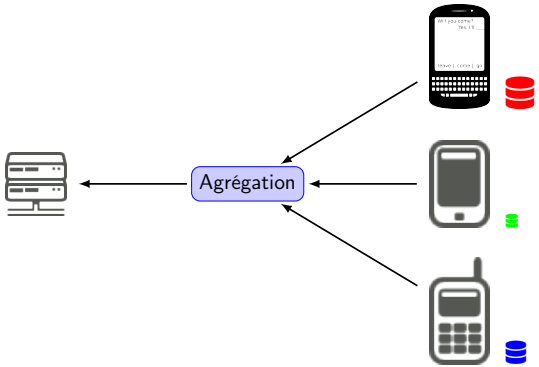


La recherche en Apprentissage Fédéré – Architecture

Hybrid

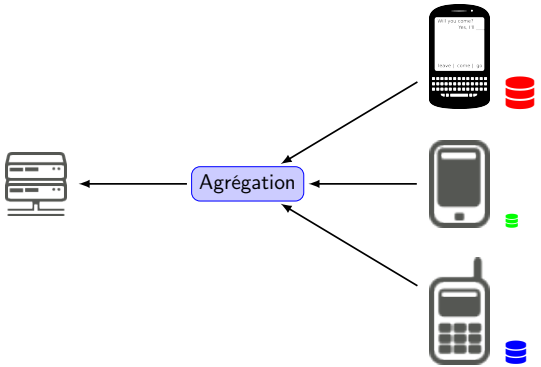


La recherche en Apprentissage Fédéré

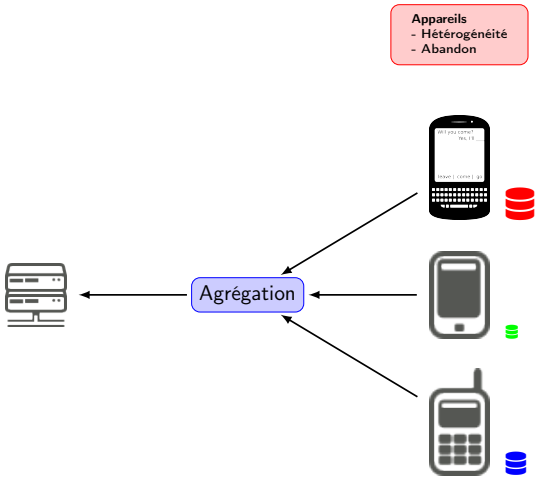


La recherche en Apprentissage Fédéré

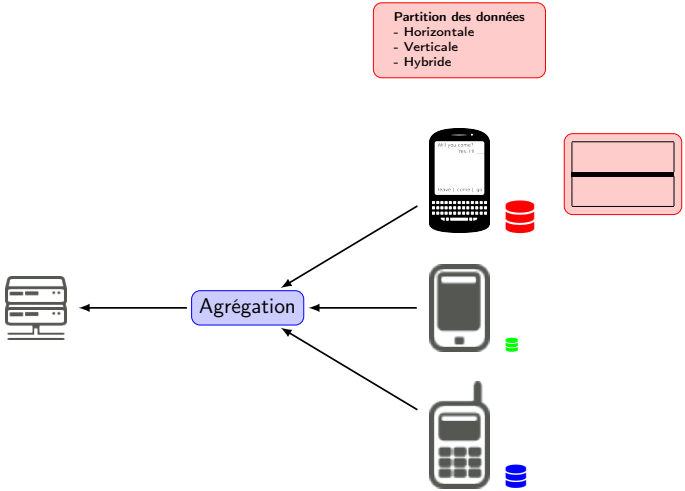
Participants
- Coopératifs
- Compétitifs
- Incitation à participer



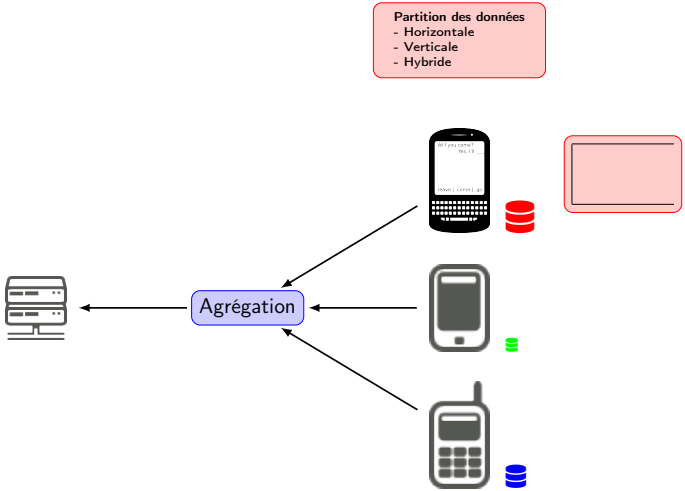
La recherche en Apprentissage Fédéré



La recherche en Apprentissage Fédéré

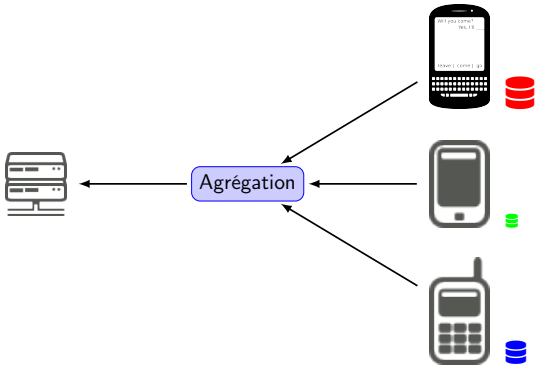


La recherche en Apprentissage Fédéré

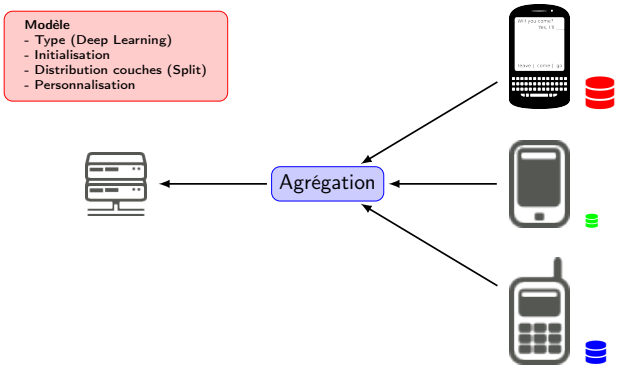


La recherche en Apprentissage Fédéré

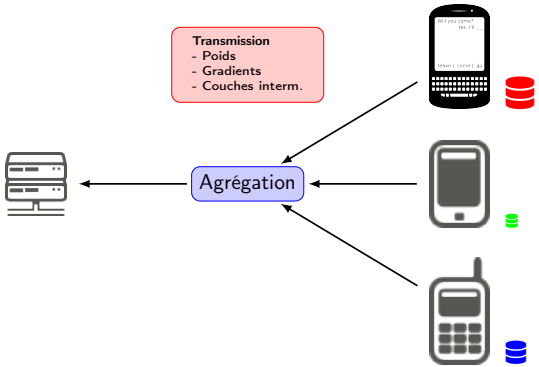
Distribution des données (non/IID)
- non/balancé
- in/dépendant
- décalage labels/attributs. . .



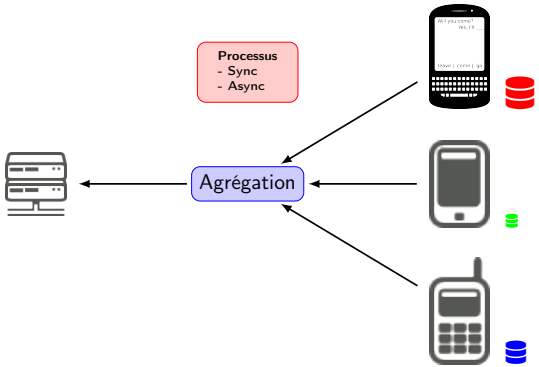
La recherche en Apprentissage Fédéré



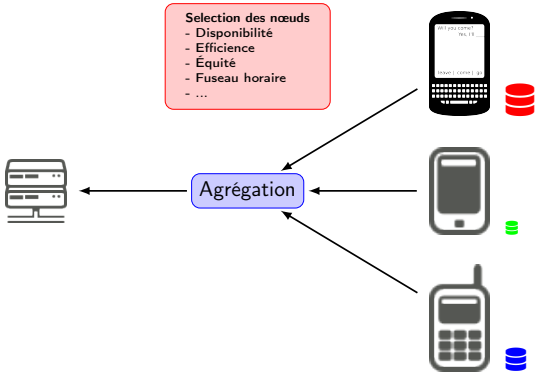
La recherche en Apprentissage Fédéré



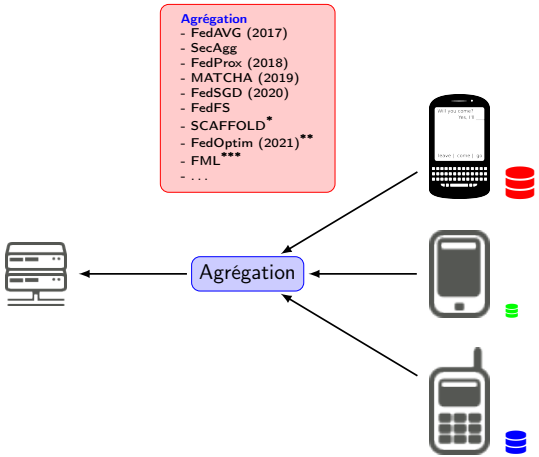
La recherche en Apprentissage Fédéré



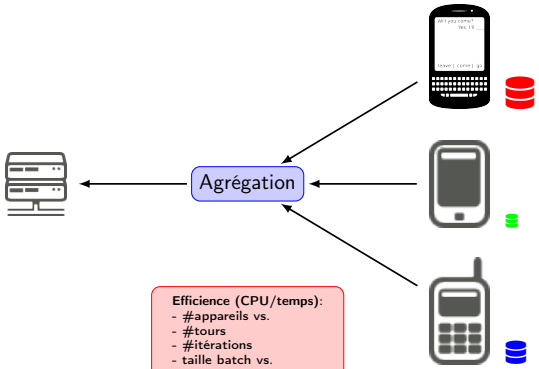
La recherche en Apprentissage Fédéré



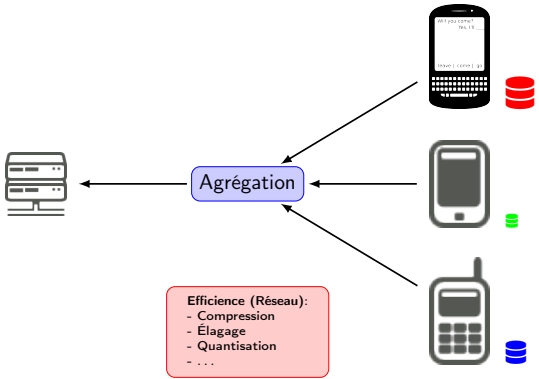
La recherche en Apprentissage Fédéré



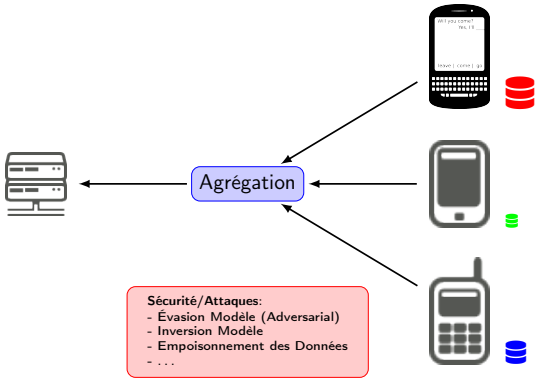
La recherche en Apprentissage Fédéré



La recherche en Apprentissage Fédéré



La recherche en Apprentissage Fédéré



Thèse Futur&Ruptures avec IMT-Atlantique & TSE

(Lucas Grativol)

Contexte

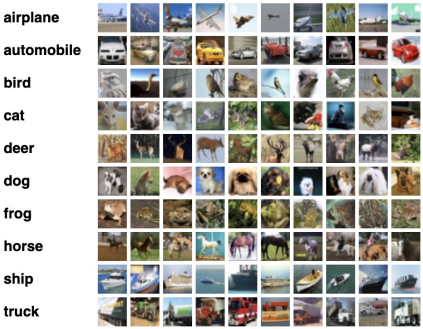
- Apprentissage Fédéré

Objectifs

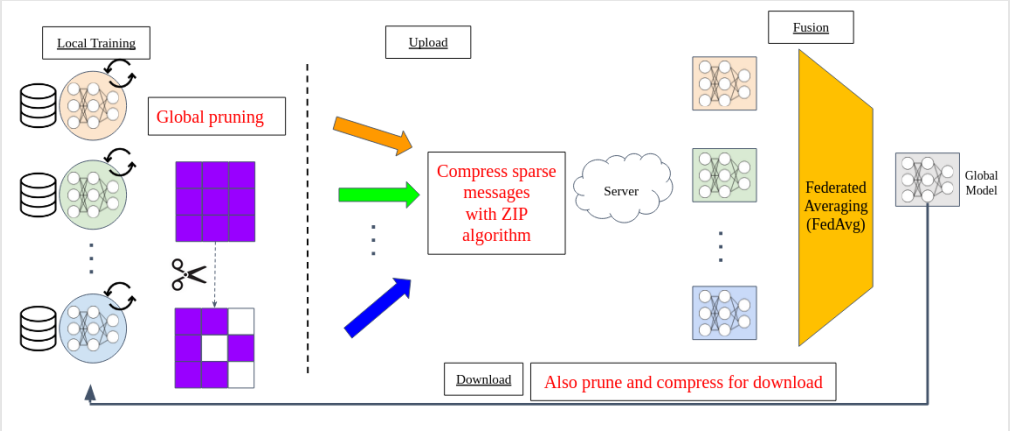
- Réduire les coûts en communication
- Réduire les coûts en calculs
- Gérer l'hétérogénéité

Contrainte

- Sans perdre en performances



Approche 1 : Élagage (“Pruning”)



Approche 1 : Élagage (“Pruning”)

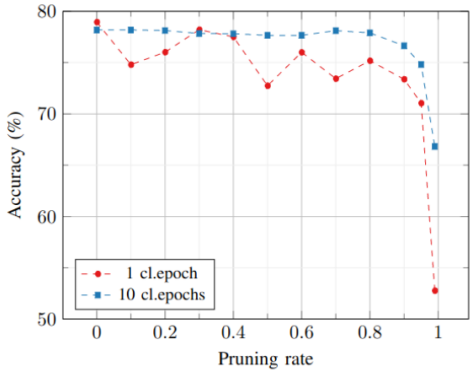
Model : ResNet - 12

Model size : 780K parameters with 2.97 MB

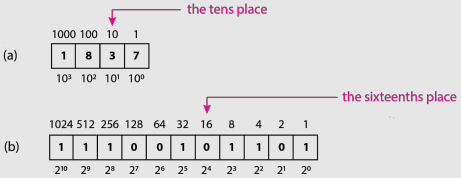
Setup : 100 rounds of communication, with 10 clients, where 40 % participate in each training round.

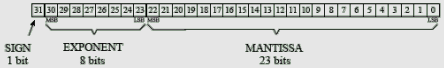
Varying the number of local epochs : 1 and 10 per round.

Tasks : Image classification on CIFAR 10 dataset.



Approche 2 : Quantisation (src)





Example 1

0 00000111 11000000000000000000000

$+1.75 \times 2^{(7-127)} = +1.316554 \times 10^{-36}$

Example 2

1 10000001 01100000000000000000000

$-1.375 \times 2^{(129-127)} = -5.500000$

Approche 2 : Quantisation (src)

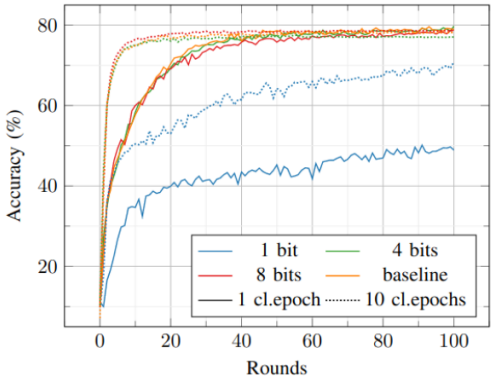
Model : ResNet - 12

Model size : 780K parameters with 2.97 MB

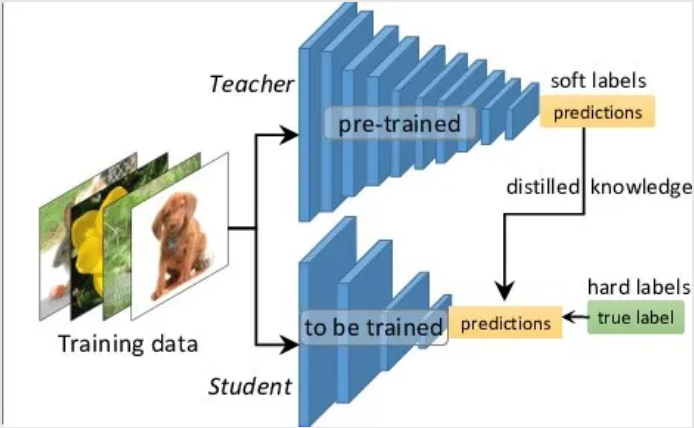
Setup : 100 rounds of communication, with 10 clients, where 40 % participate in each training round.

Varying the number of local epochs : 1 and 10 per round.

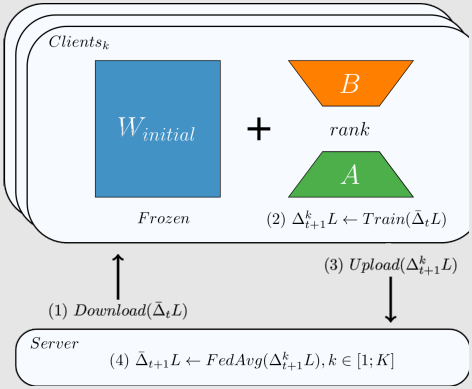
Tasks : Image classification on CIFAR 10 dataset.



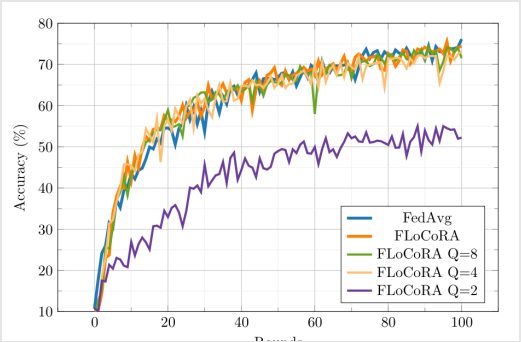
Approche 3 : Distillation (src)



Approche 4 : Low Rank Adaptation (LoRA)



Approche 4 : Low Rank Adaptation (LoRA)



Method	Total Params	Trained Params	% of Trained Params
FedAvg	**1.23M**	**1.23M**	**100**
FLoCoRA ($r = 8$)	1.30M	69.45K	5.35
FLoCoRA ($r = 16$)	1.36M	131.92K	9.70
FLoCoRA ($r = 32$)	1.48M	256.84K	17.30
FLoCoRA ($r = 64$)	1.73M	506.70K	29.22
FLoCoRA ($r = 128$)	2.23M	1.00 M	45.05

Thèse+Stage Master Cyber-Sécurité

(Ons Aouedi, Ahlem Harhad)

Contexte

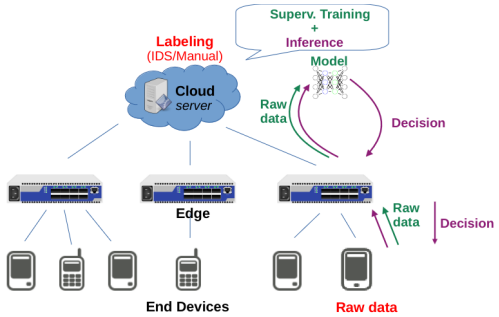
- Cyber-Sécurité en cœur de réseau

Objectifs

- Détecter attaque plus proche/tôt
- Réduire les coûts en communication
- Réduire les coûts en calculs

Contraintes

- Sans perdre en performances
- Le plus automatisé possible (~~labels~~)



Approche Semi-Supervisée & Fédéré / TinyML

4. Résultats

FLuIDS a été testé sur 2 jeux de données classiques de la détection d'intrusions, notamment IIoT : UNSW-NB15 et SCADA Gas Pipeline. Les résultats suivants portent sur le premier. Il a pu être montré que :

- FLuIDS **exploite les données non-étiquetées** : augmenter leur volume (à nb. étiquetées constant), augmente l'accuracy.
- FLuIDS **réduit** de 20% à 99% **le volume de communications**, en envoyant des mise-à-jour de modèles plutôt que des données brutes.
- FLuIDS est **au moins aussi performant** (F1-score) que les modèles **centralisés supervisés**.
- FLuIDS est **plus performant** quand il est entraîné **en mode fédéré** qu'en mode centralisé : il est peu sensible à une distribution non-IID des données.
- FLuIDS nécessite **plus temps pour converger** qu'un modèle centralisé, mais dans une limite ***raisonnable***.

Stage Master+Projets Recherche

(A. Gaffarov, Q. Rey, I. Perez)

Contexte

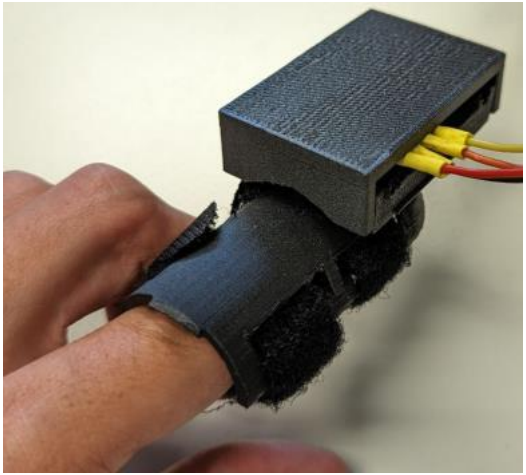
- Médecine "augmentée"

Objectifs

- Interpréter les mouvements...
- (... pour afficher de l'aide)

Contraintes

- Sur (tout) petit appareil
- En autonomie
- Sans perdre en performance



Approches: Transfert Learning & Few-Shot Learning

Data processing

Evaluation method

Person dependant: Test on a random sample, including individuals from the training set

Person independent: Test on a new individual not included in the training set

Models and results

	Characteristics		
LSTM	Exploit temporal component	90%	**80%**
CNN	Exploit patterns in the image	80%	70%

Merci de votre attention

MINISTÈRE DE L'ÉCONOMIE, DES FINANCES ET DE LA SOUVERAINETÉ INDUSTRIELLE ET NUMÉRIQUE

Institut Mines-Télécom

INSPIRING INNOVATION SINCE 1816

#+EXPORT$_{\text{LATEX}}$ |llll|%
% &~&~&\
&~&~&\
% &~&~&\
&~&~&\
% % } {|lll!{}|}%
% &~&~&\
&~&~&\
&~&~&\
&~&~&\
% % }
||# + *END*$_{\text{EXPORT}}$