

Quelques problèmes d'apprentissage dans le cadre de l'Internet des objets

Massinissa Hamidi

sous la direction de Aomar Osmani

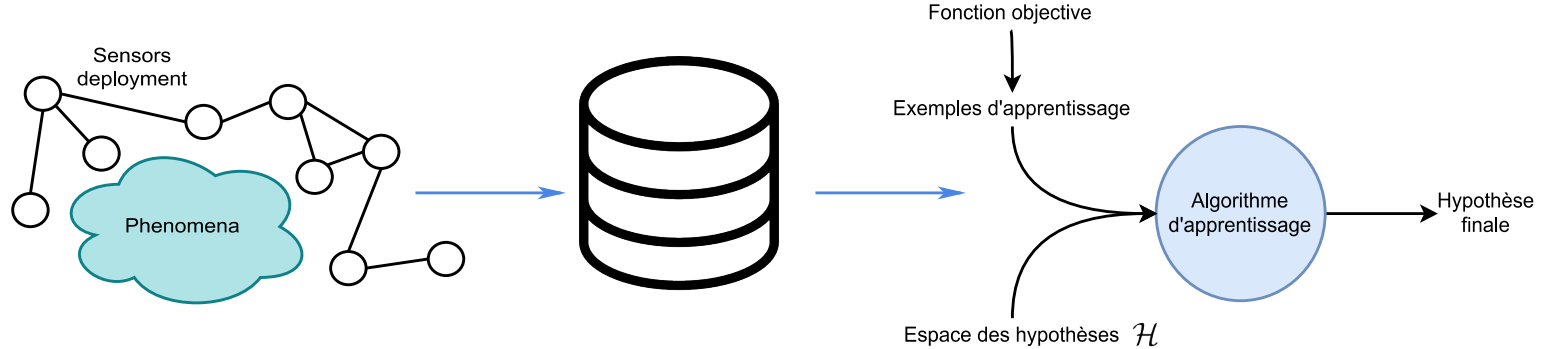
LIPN-UMR CNRS 7030, Univ. Sorbonne Paris Nord



Contexte et motivations



Apprentissage dans l'Internet des objets

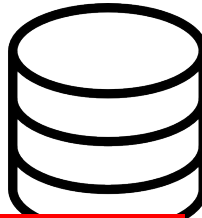
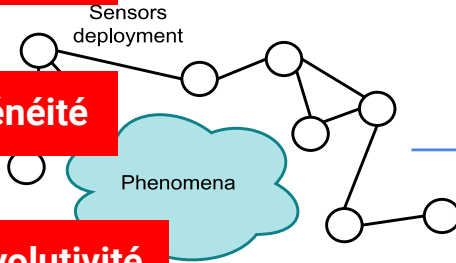


Apprentissage dans l'Internet des objets

Perspectives multiples

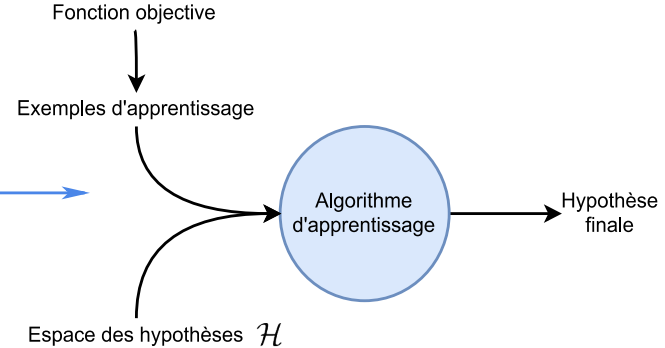
Hétérogénéité

Évolutivité



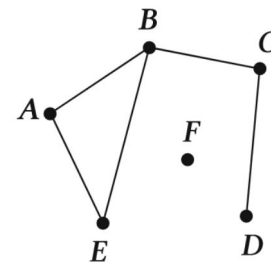
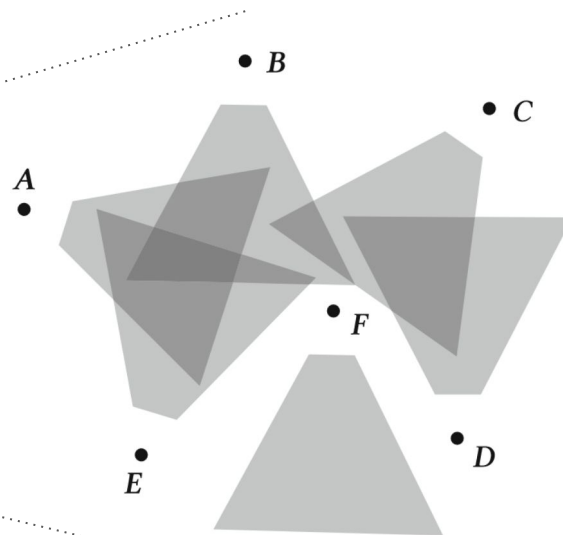
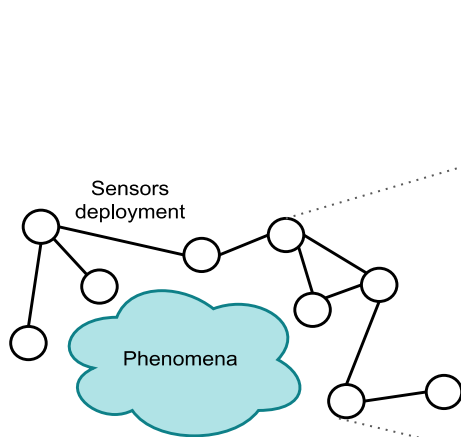
FIGÉ

Biais inductifs



Spécificités des déploiements

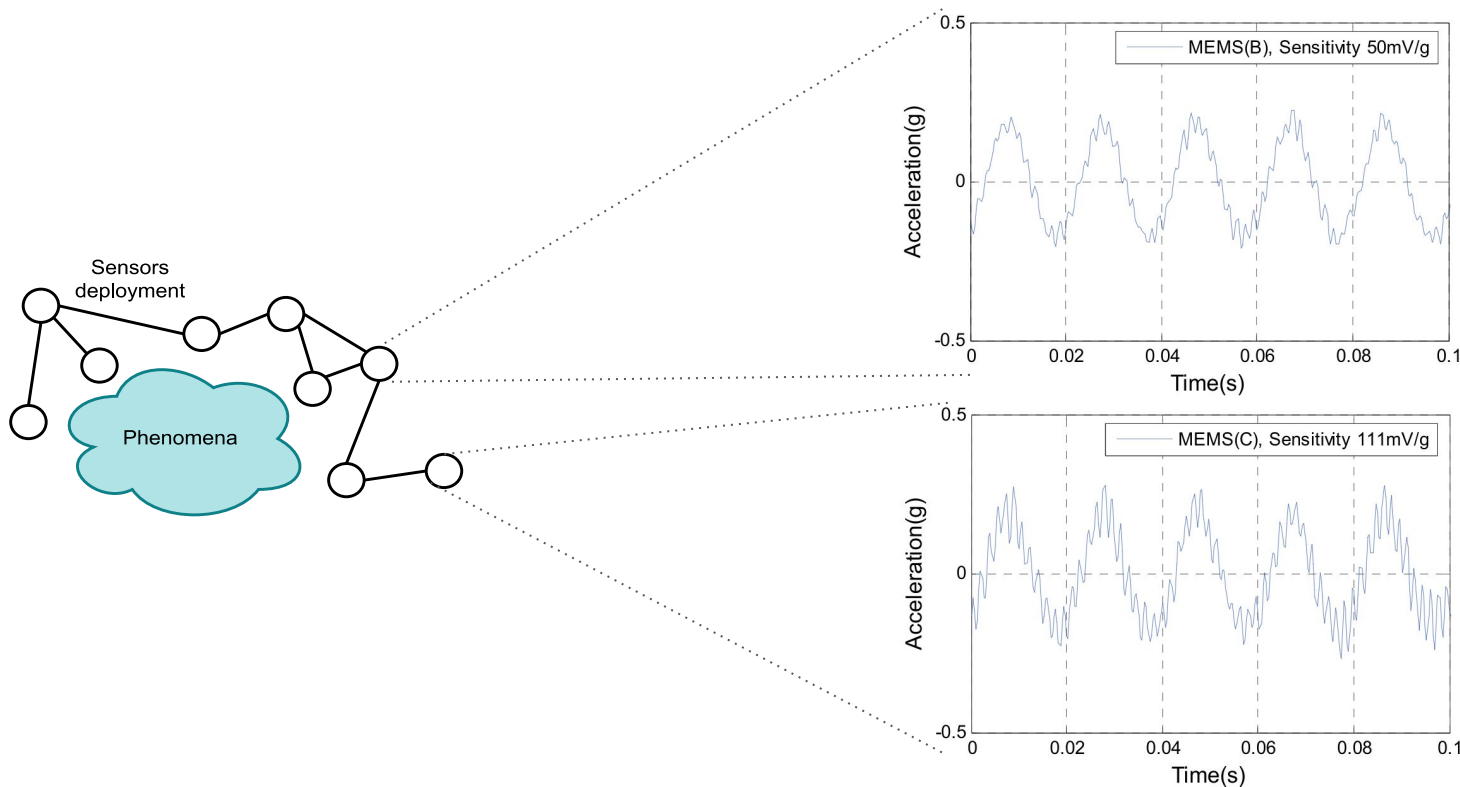
Topologies des déploiements



- Plusieurs perspectives : redondantes, complémentaires, contradictoires, etc. ;
- Privacy constraints according to the perspectives and their evolution over time ;

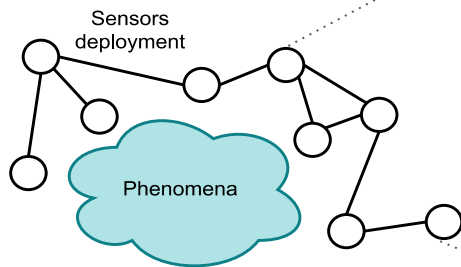
Spécificités des déploiements

Hétérogénéité des sources de données : impact de la sensibilité des capteurs sur les mesures

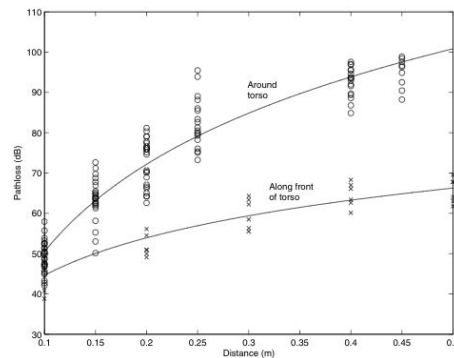
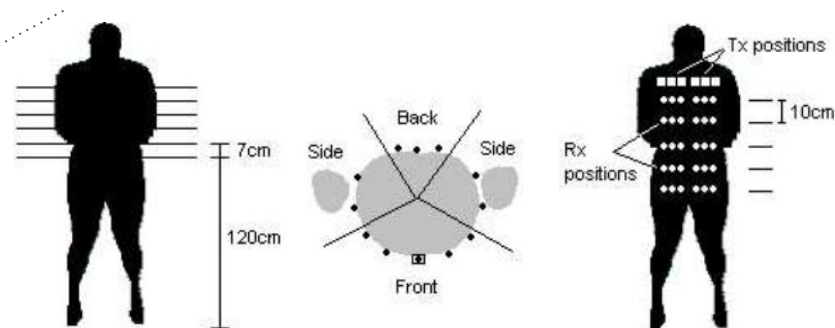


Spécificités des déploiements

Transmissions: Path loss; Shadowing



Traitements circonscrits au niveau dans lequel ces problématiques apparaissent, e.g. energy-aware et temperature-aware routing protocols, propagation loss models, etc.

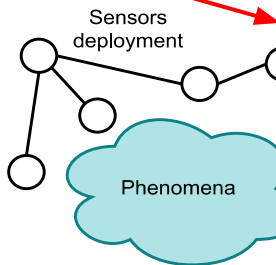


Spécificités des déploiements

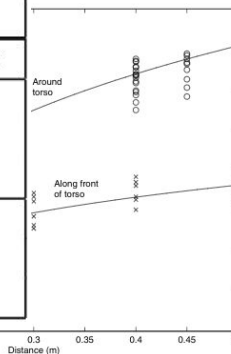
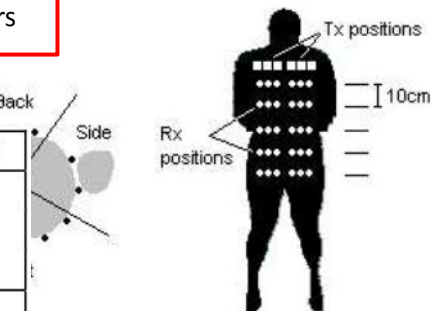
Transmissions: Path loss; Shadowing;

Position des transmetteur/récepteurs

Environnement

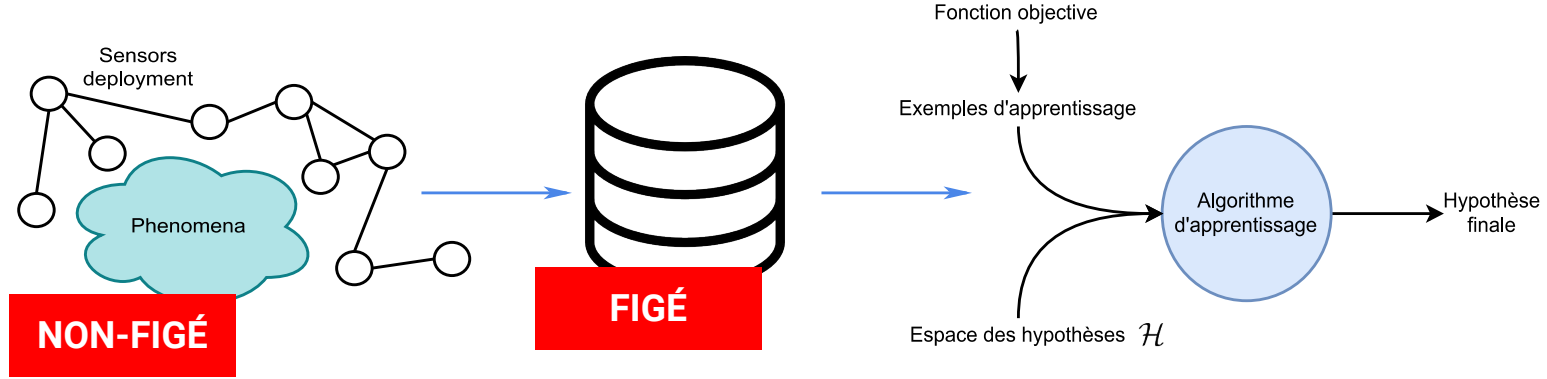


TX on Hip / RX on		Chest	Thigh	R Wrist	R Foot
Anech.	Still	0.61	0.41	0.95	0.41
	Walking	2.15	5.27	4.31	4.97
	Running	2.39	2.47	3.49	2.69
Indoor	Still	0.60	0.24	0.26	0.24
	Walking	1.52	3.27	2.66	2.57
	Running	2.00	1.98	2.37	1.80
TX on L Ear / RX on		R Ear	Hip	R Wrist	R Foot
Anech.	Still	0.23	0.45	1.15	0.56
	Walking	0.54	2.07	3.35	4.21
	Running	0.52	2.82	2.98	1.57
Indoor	Still	0.21	0.28	0.31	0.30
	Walking	0.90	2.20	1.80	2.02
	Running	0.70	2.19	1.88	1.24

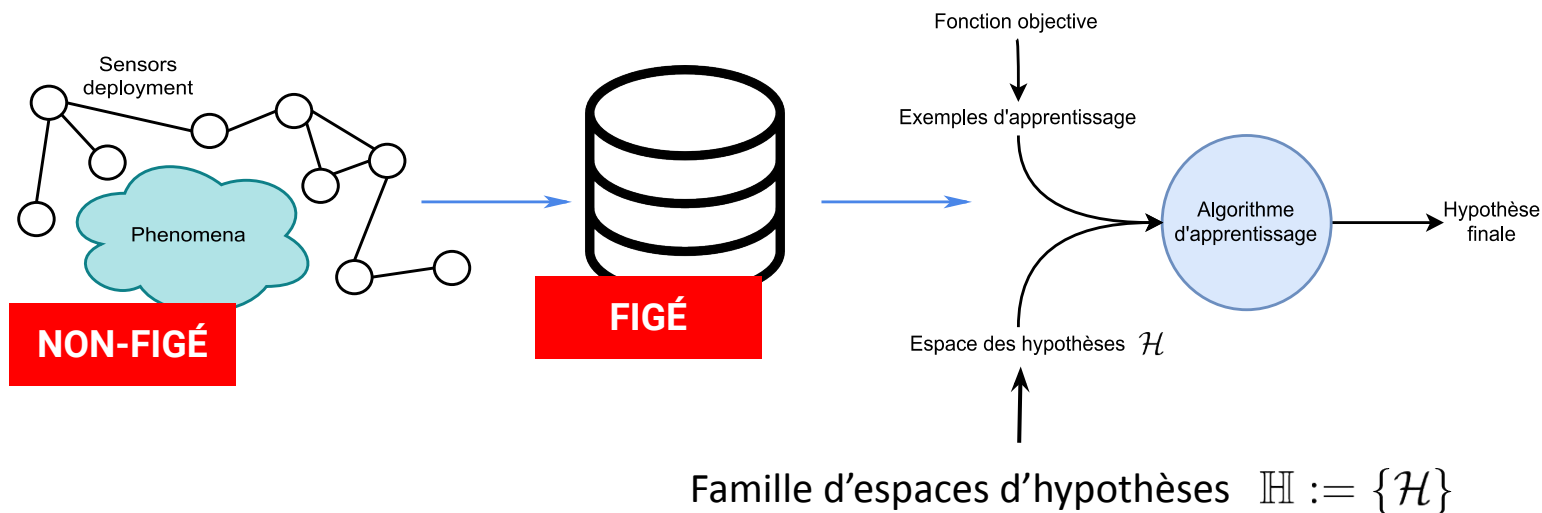


Type d'activité

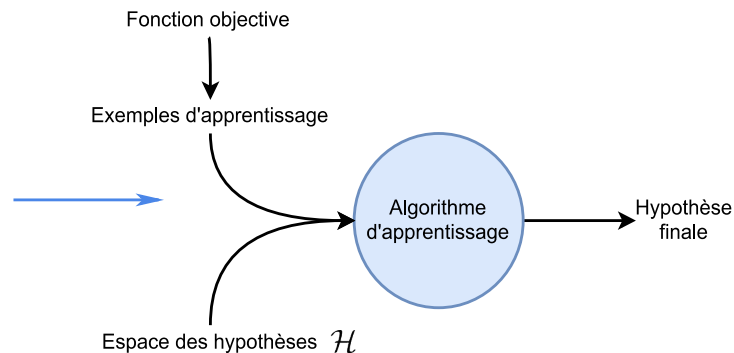
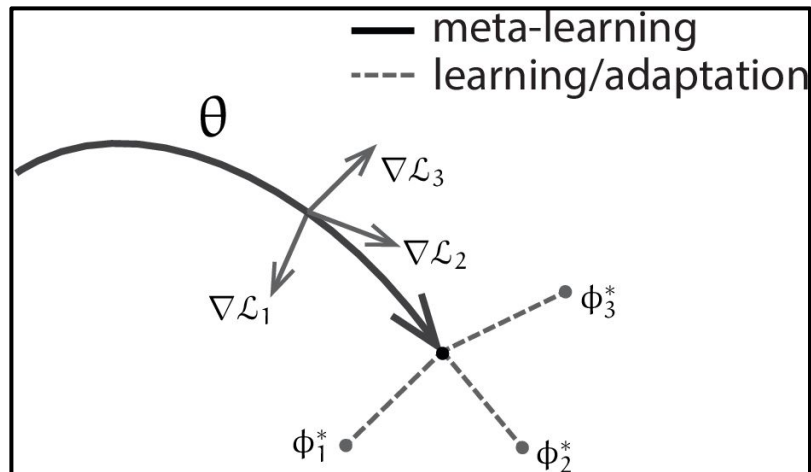
Apprentissage dans l'Internet des objets



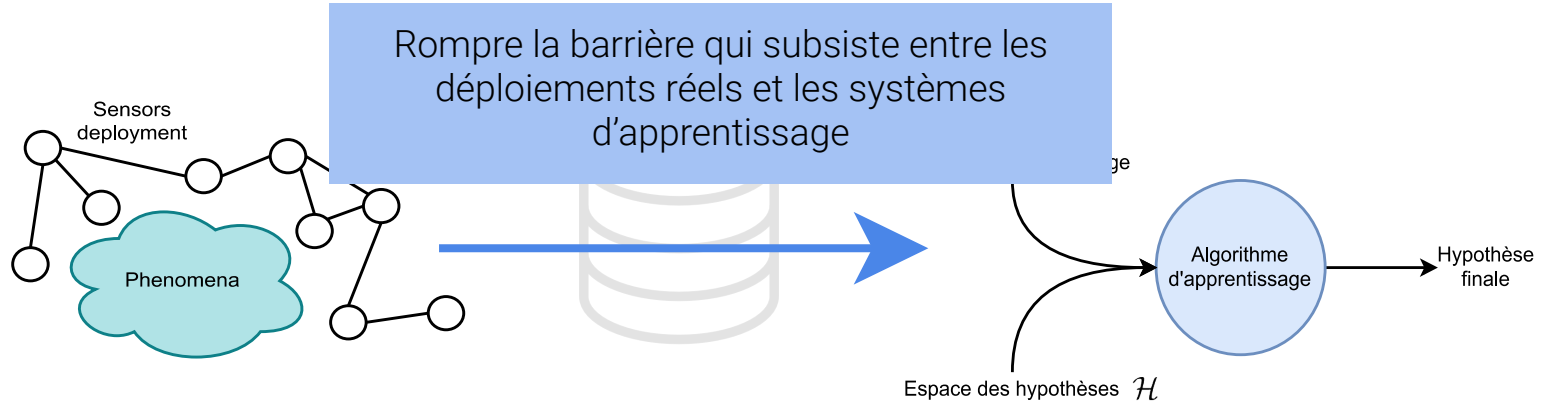
Apprentissage dans l'Internet des objets



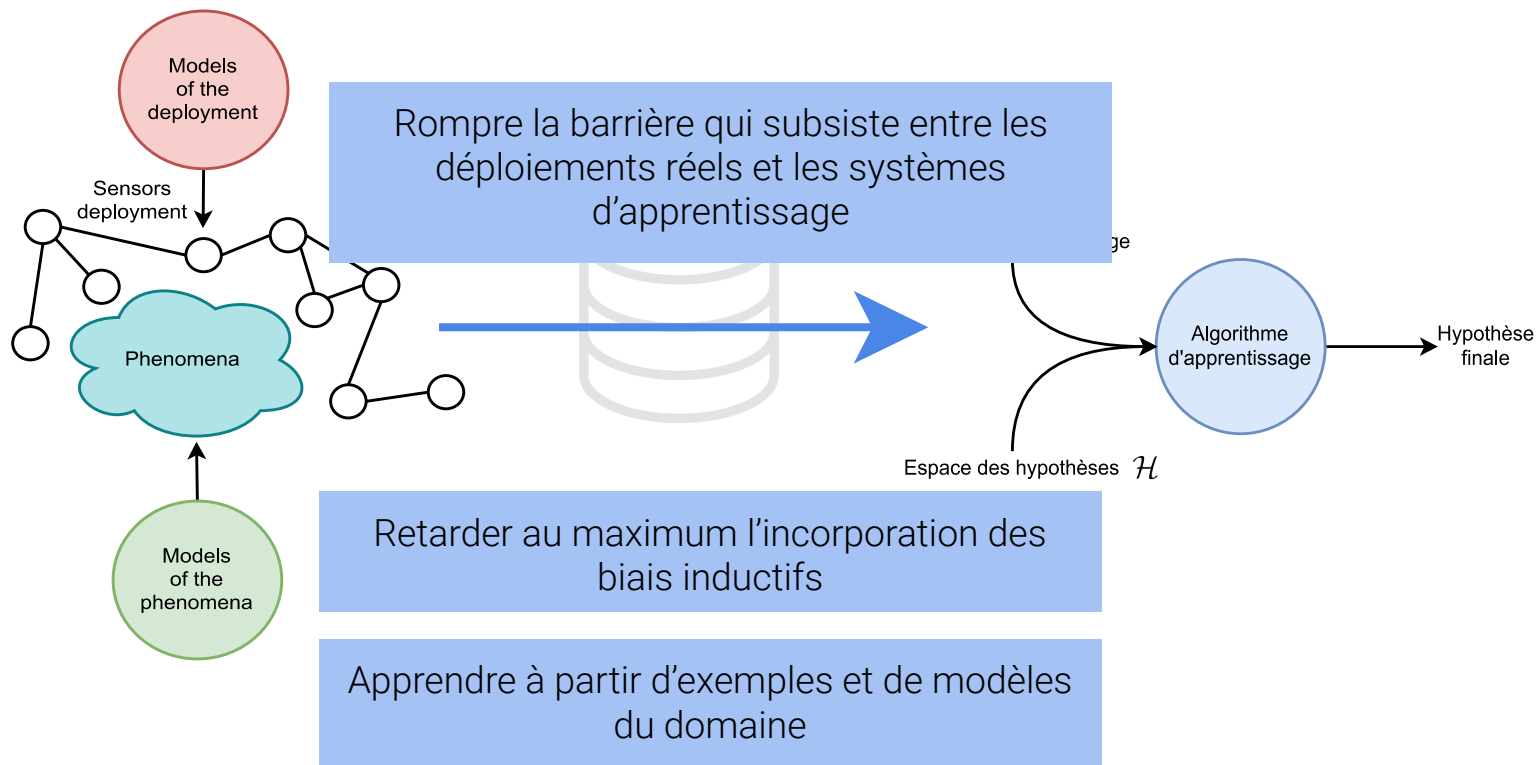
Apprentissage dans l'Internet des objets



Apprentissage dans l'Internet des objets

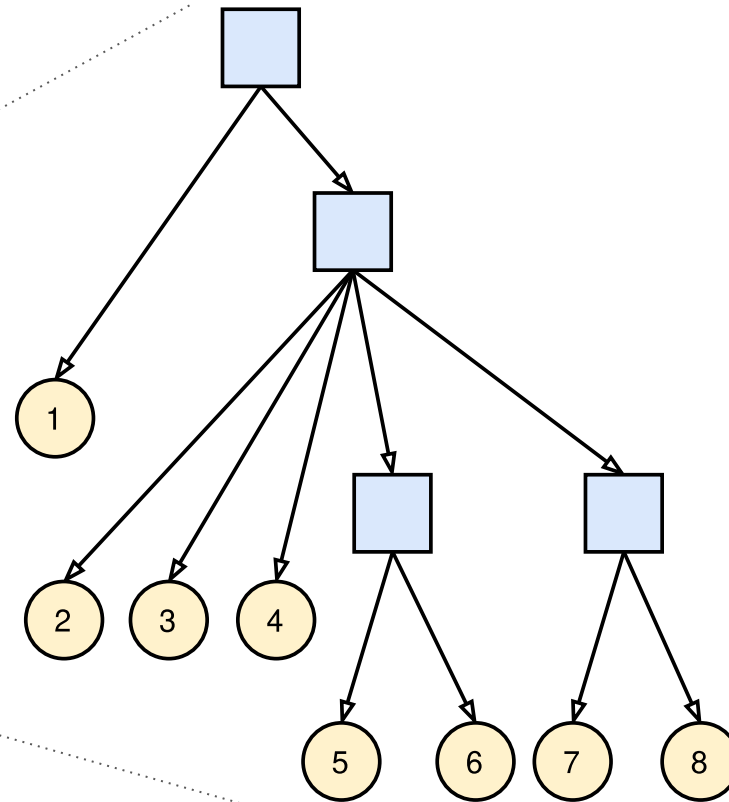
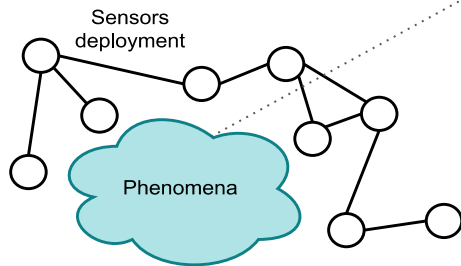


Apprentissage dans l'Internet des objets



Spécificités des phénomènes

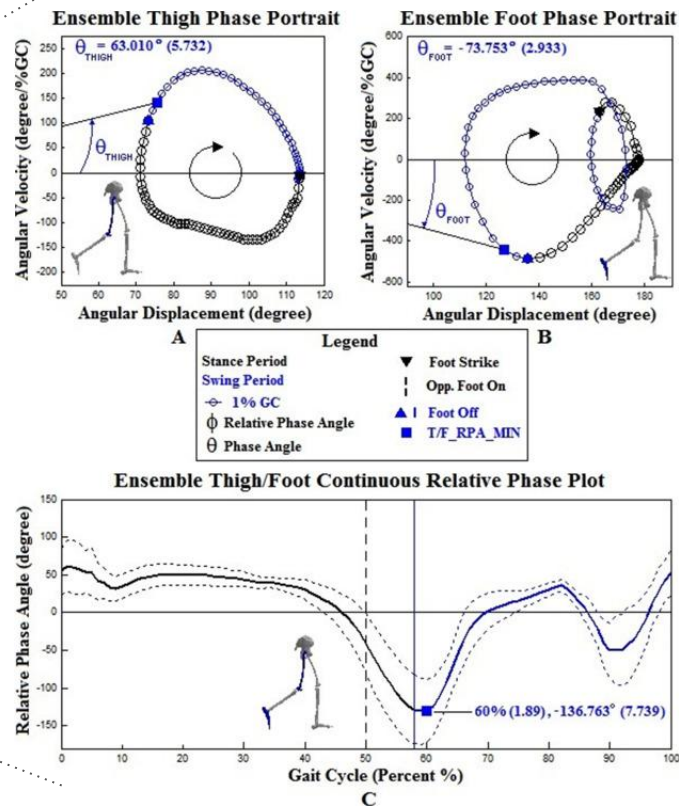
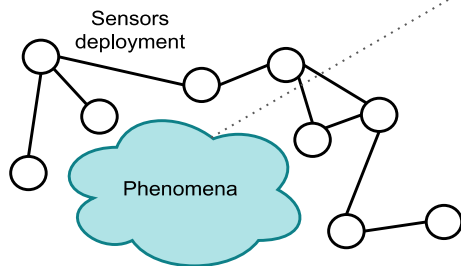
Dépendance entre concepts et leurs structuration



Spécificités des phénomènes

Modèles biomécaniques des mouvements

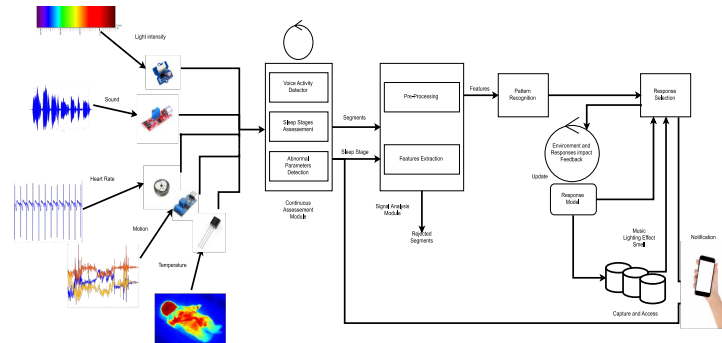
Modèles biomécaniques des mouvements
qui décrivent la dynamique des mouvements
dans le temps



Applications Explorées

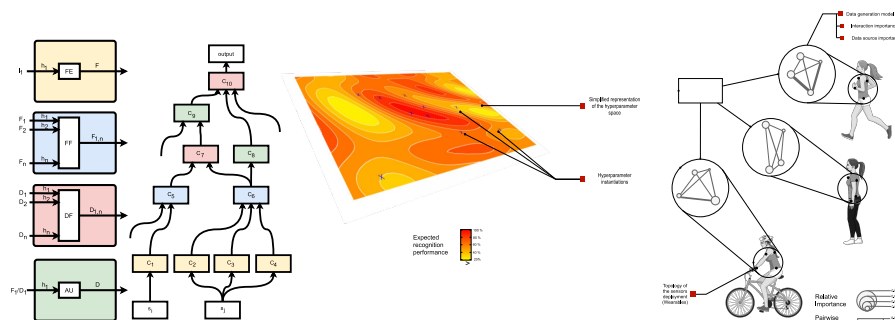
Applications Explorées

- Reconnaissance des situations d'inconfort chez les nourrissons ;
- Reconnaissance d'activités humaines ;
- Suivi du phénomène vibratoire des turbocompresseurs ;
- Synthèse de nouveaux matériaux en industrie



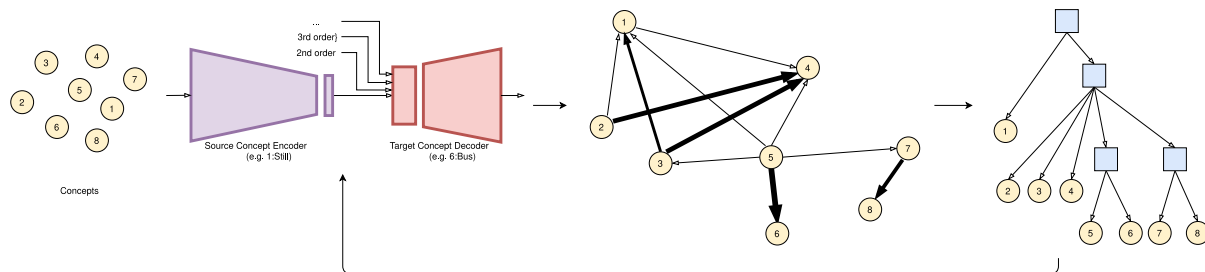
Applications Explorées

- Reconnaissance des situations d'inconfort chez les nourrissons ;
- Reconnaissance d'activités humaines ;
- Suivi du phénomène vibratoire des turbocompresseurs ;
- Synthèse de nouveaux matériaux en industrie



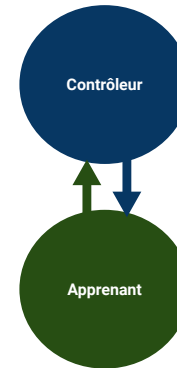
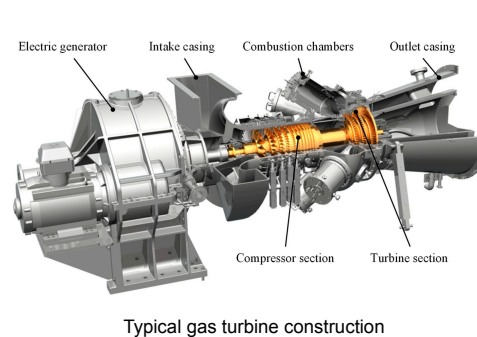
Applications Explorées

- Reconnaissance des situations d'inconfort chez les nourrissons ;
- Reconnaissance d'activités humaines ;
- Suivi du phénomène vibratoire des turbocompresseurs ;
- Synthèse de nouveaux matériaux en industrie



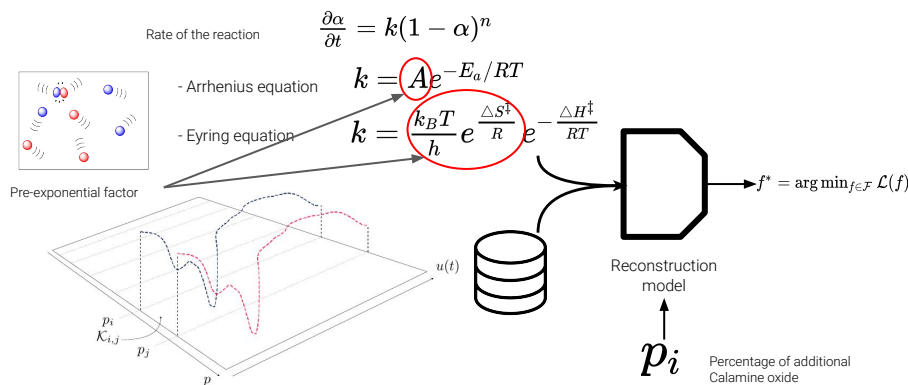
Applications Explorées

- Reconnaissance des situations d'inconfort chez les nourrissons ;
- Reconnaissance d'activités humaines ;
- Suivi du phénomène vibratoire des turbocompresseurs ;
- Synthèse de nouveaux matériaux en industrie



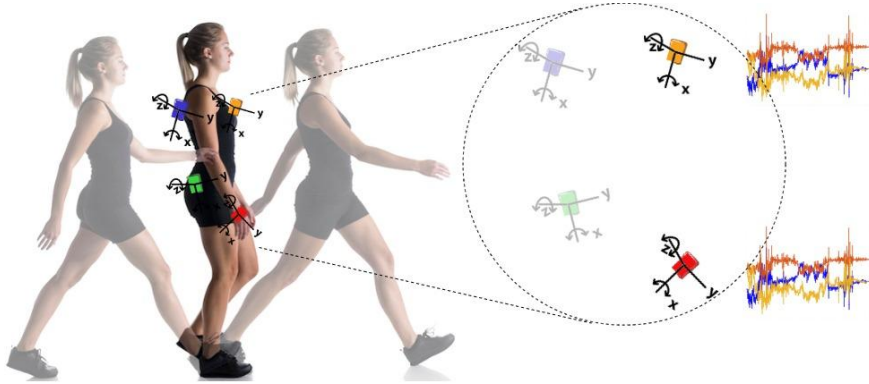
Applications Explorées

- Reconnaissance des situations d'inconfort chez les nourrissons ;
- Reconnaissance d'activités humaines ;
- Suivi du phénomène vibratoire des turbocompresseurs ;
- Synthèse de nouveaux matériaux en industrie

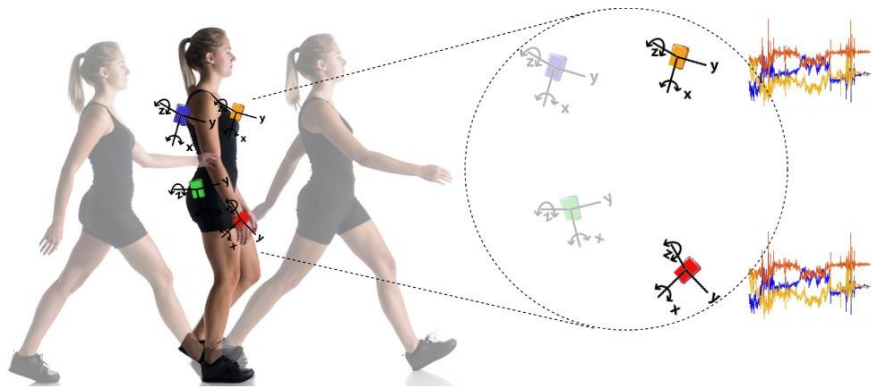


Abstraction de la position des capteurs

Reconnaissance d'activités à partir de déploiements

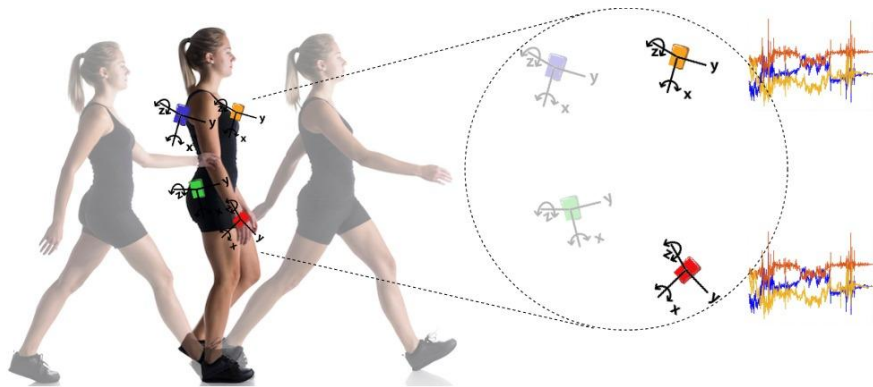


Reconnaissance d'activités à partir de déploiements



on considère des configuration où
une collection \mathcal{S} de M capteurs,
notée $\{s_1, \dots, s_M\}$, sont positionnées
sur l'objet d'intérêt aux positions
 $\{p_1, \dots, p_M\}$.

Reconnaissance d'activités à partir de déploiements

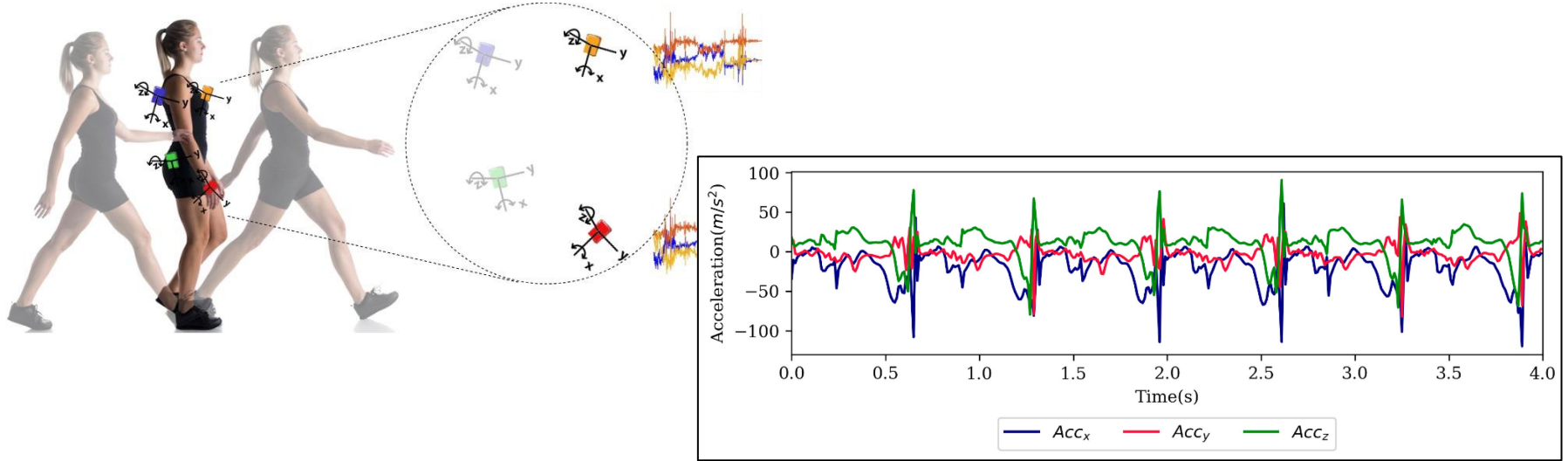


on considère des configuration où une collection \mathcal{S} de M capteurs, notée $\{s_1, \dots, s_M\}$, sont positionnées sur l'objet d'intérêt aux positions $\{p_1, \dots, p_M\}$.

Chaque capteur s_i génère un flux $\mathbf{x}^i = (x_1^i, x_2^i, \dots)$ d'observations d'une certaine modalité comme l'accélération ou la gravité.

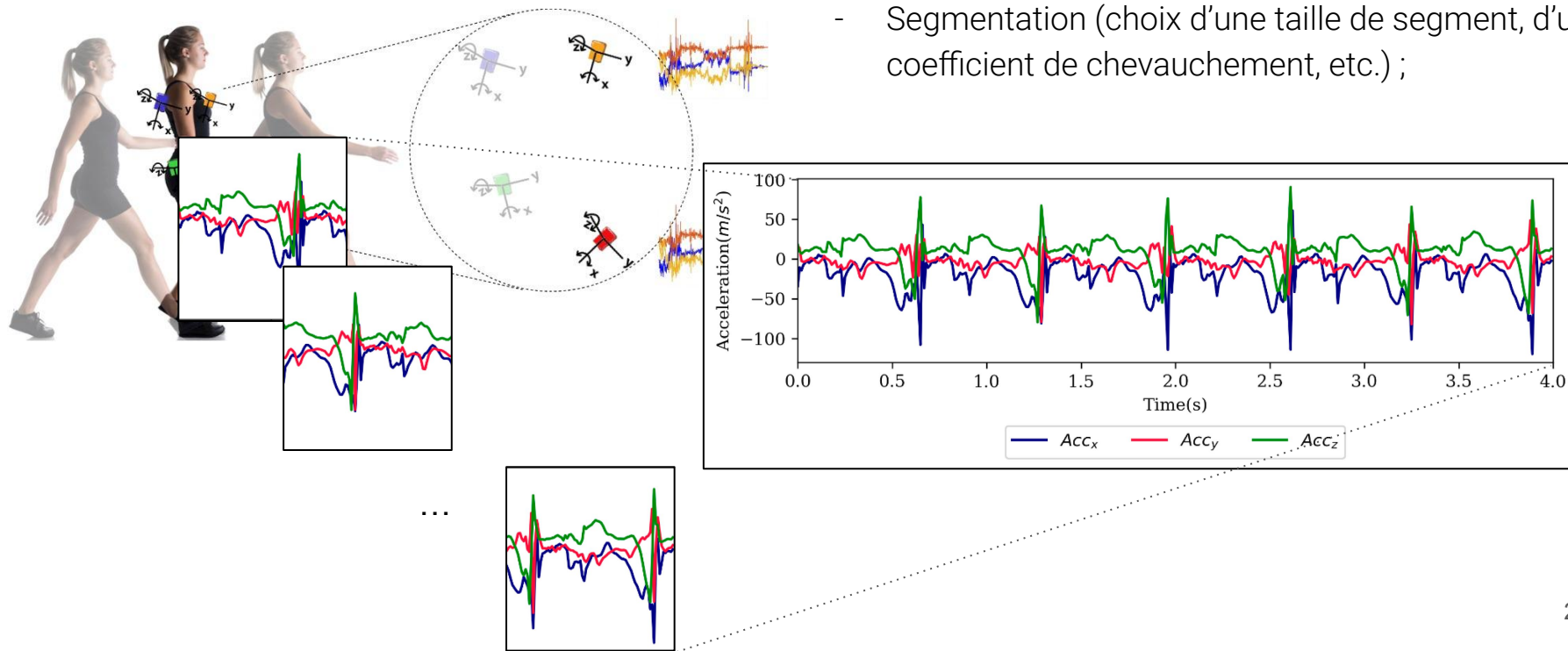
Reconnaissance d'activités à partir de déploiements

- Acc, Gyr, Mag, etc. ;
- Séquences temporelles ;

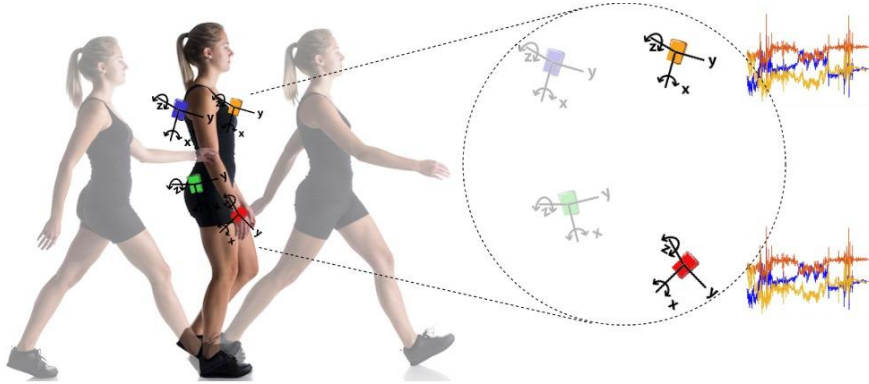


Reconnaissance d'activités à partir de déploiements

- Acc, Gyr, Mag, etc. ;
- Séquences temporelles ;
- Segmentation (choix d'une taille de segment, d'un coefficient de chevauchement, etc.) ;

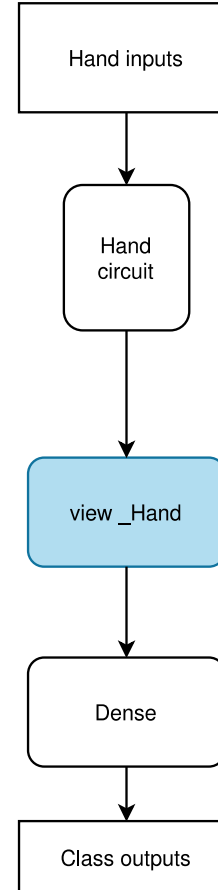
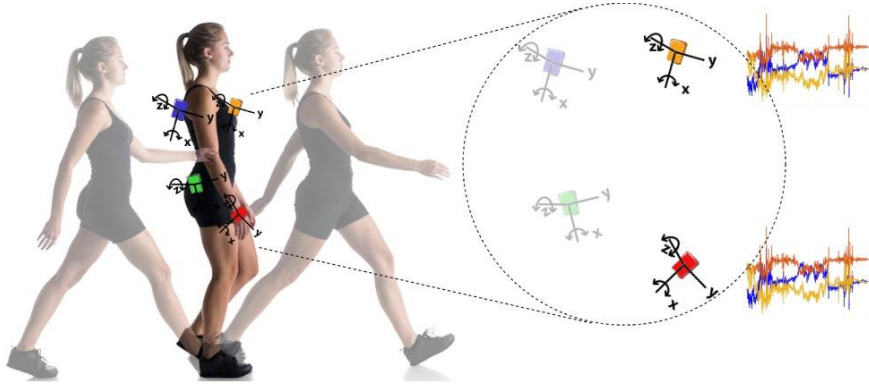


Reconnaissance d'activités à partir de déploiements

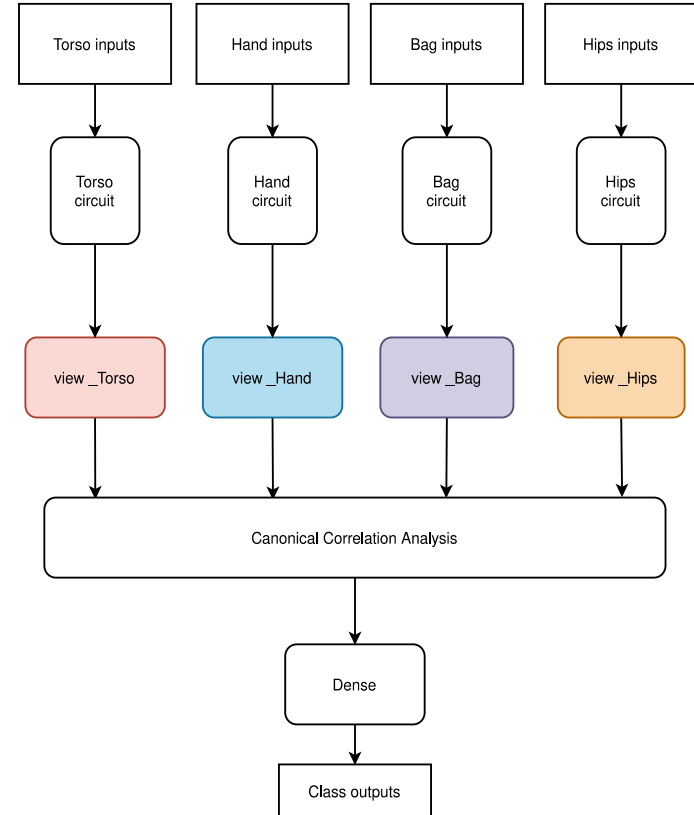
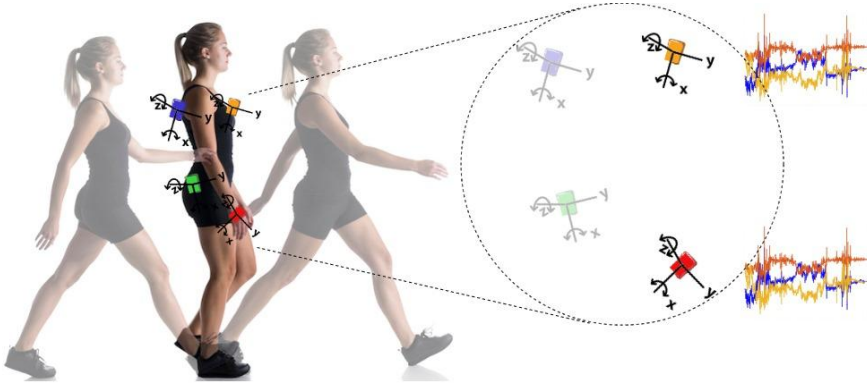


- Acc, Gyr, Mag, etc. ;
- Séquences temporelles ;
- Segmentation (choix d'une taille de segment, d'un coefficient de chevauchement, etc.) ;
- Généralement, on se ramène à un problème de classification de séquences (avec des réseaux de neurones par exemple chargés d'extraire des caractéristiques pertinentes du signal, etc. ;
- On utilise des LSTM afin de modéliser les dépendances temporelles du signal ;

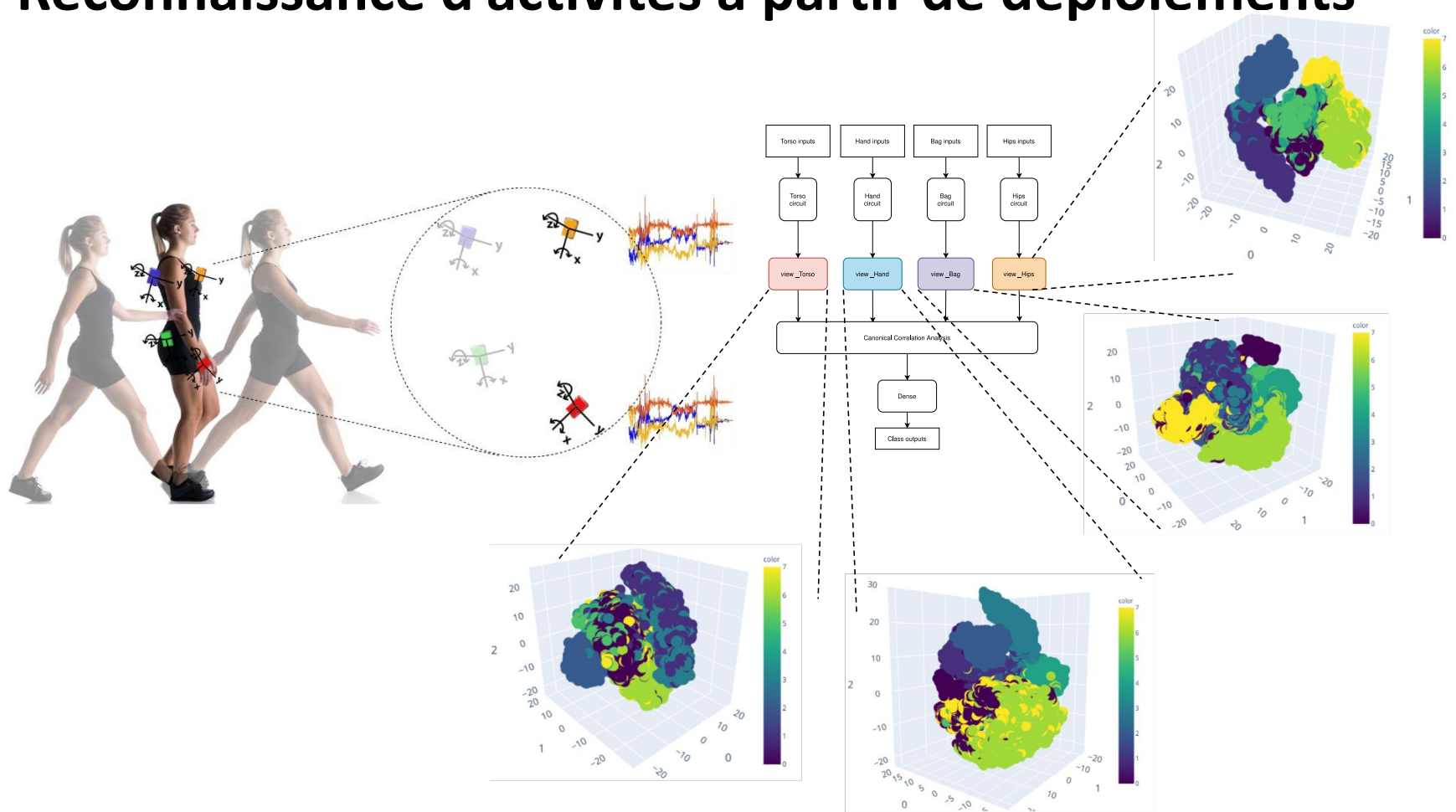
Reconnaissance d'activités à partir de déploiements



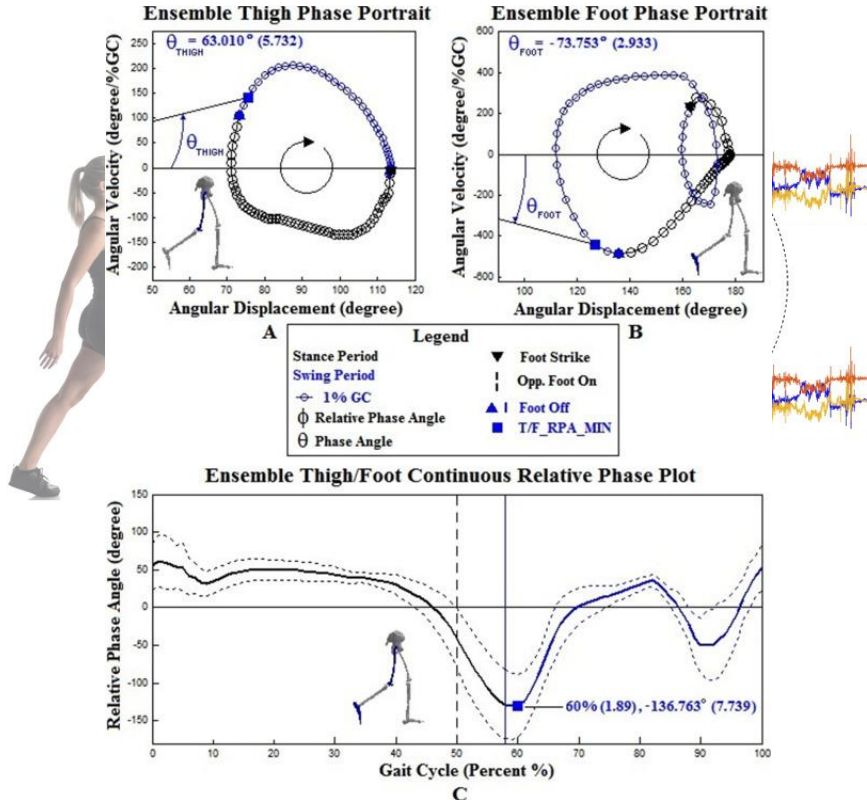
Reconnaissance d'activités à partir de déploiements



Reconnaissance d'activités à partir de déploiements



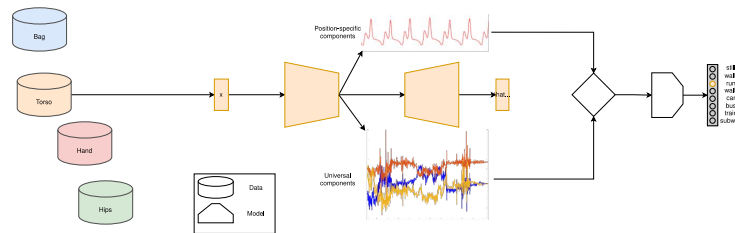
Reconnaissance d'activités à partir de déploiements



- Mouvements translationnel et circulaire ;
- Discriminer entre marcher et courir peut se baser simplement sur les composantes translationnelles vu qu'il y a un déplacement plus important par rapport à l'environnement ;
- Discriminer entre courir et être à vélo peut se faire en utilisant les composantes cycliques vu que la main reste relativement statique à vélo.

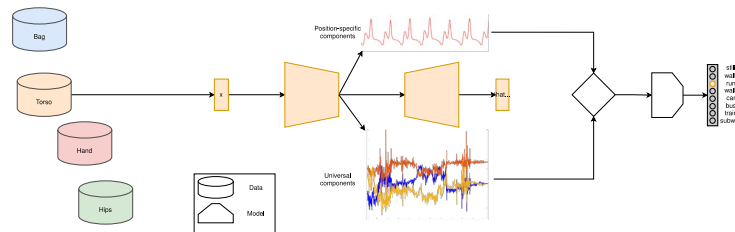
Abstraction de la Position des Capteurs

Grâce à cette transformation chaque capteur $s \in \mathcal{S}$ démêlera les flux de données entrelacés entre le composant local et universel \mathbf{x} en les projetant dans deux représentations distinctes z_A et z_P .



Abstraction de la Position des Capteurs

Grâce à cette transformation chaque capteur $s \in \mathcal{S}$ démêlera les flux de données entrelacés entre le composant local et universel \mathbf{x} en les projetant dans deux représentations distinctes \mathcal{Z}_A et \mathcal{Z}_P .



Les représentations \mathcal{Z}_P ne seront utilisées qu'au niveau de l'apprenant local qui les a générées. Les représentations \mathcal{Z}_A peuvent être utilisées au niveau de l'apprenant local ou partagées (mutualisées) avec les représentations de même type provenant des autres apprenants.

Approches de démêlement (*disentanglement*)

Moving MNIST

Groundtruth	7	8	5	0	3	3	5	2	5	8	9
Prediction	7	8	5	0	3	3	5	2	5	8	9
1st Component	7	8	6	4	3	6	5	2	5	8	
2nd Component	2	1	5	0	7	3	3	0	0	9	

Input

Prediction



Approches de démêlement (*disentanglement*)

Moving MNIST

Groundtruth	7	8	5	0	3	3	5
Prediction	7	8	5	0	3	3	5
1st Component	7	8	6	4	3	6	5
2nd Component	2	1	5	0	7	3	

Input

Prediction

dSprites dataset

								Data
								recons
								recons
								recons
								recons
								recons

3.1 nats 7.3 nats 15.4 nats 23.8 nats

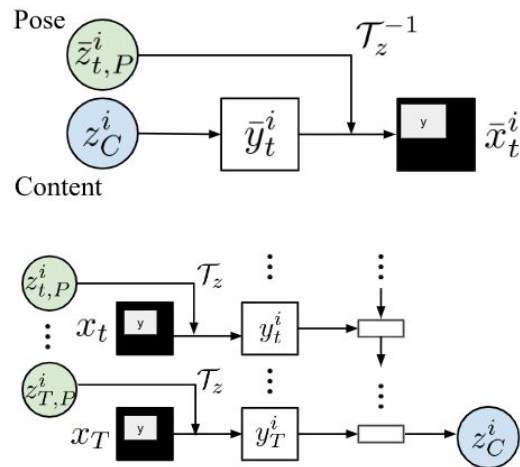
Approches de démêlement (*disentanglement*)

Moving MNIST

Groundtruth	7	8	5	0	4	3	3	5	3	2	0	5	0	8	9
Prediction	7	8	5	0	4	3	3	5	3	2	0	5	0	8	9
1st Component	7	8	6	4	3	6	5	2	5	8					
2nd Component	7	1	5	0	7	3	3	0	0	9					

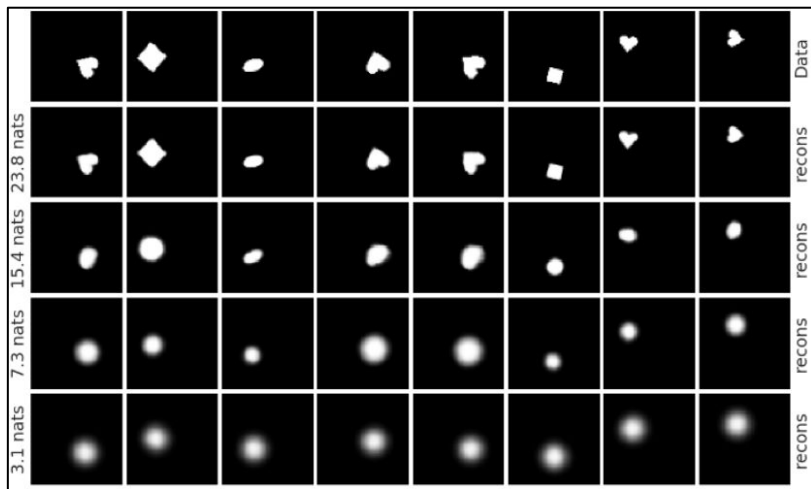
Input
 Prediction

Séparation explicite des composantes de la représentation latente :



Approches de démêlement (*disentanglement*)

dSprites dataset

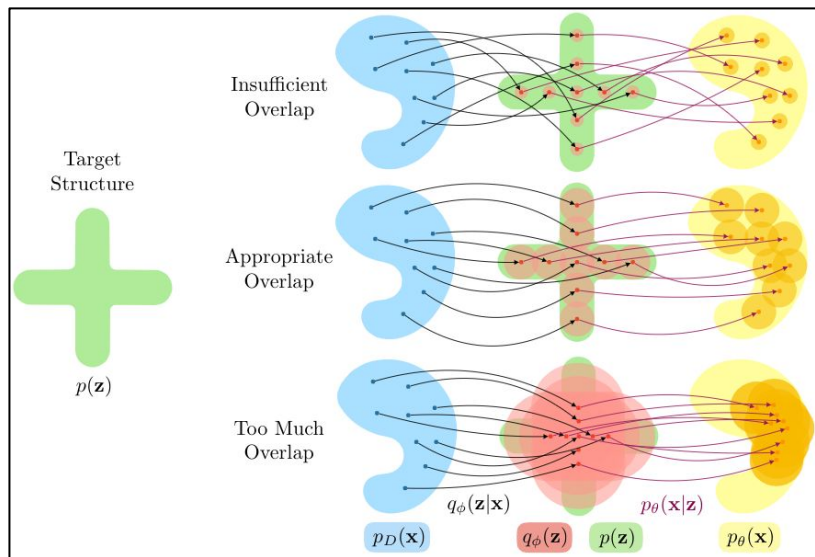


Séparation ou décomposition implicite des composantes de la représentation latente à l'aide du β -VAE

$$L(\theta, \varphi; x, z) = \mathbb{E}_{q_{\varphi}(z|x)}[\log p_{\theta}(x|z)] - \beta D_{KL}(q_{\varphi}(z|x) || p(z))$$

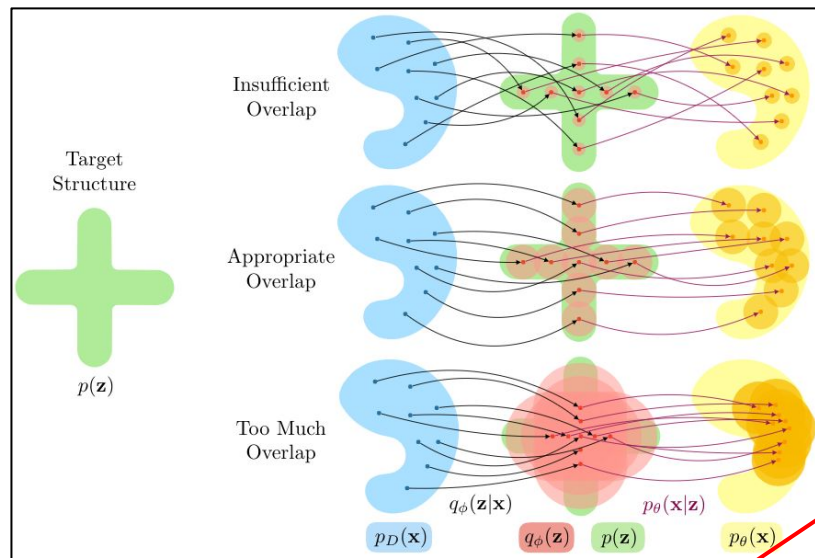
Approches de démêlement (*disentanglement*)

Imposition d'une structure particulière à l'espace latent :



$$L(\theta, \varphi; x, z) = \mathbb{E}_{q_\varphi(z|x)} [\log p_\theta(x|z)] - \beta D_{KL}(q_\varphi(z|x) || p(z)) - \alpha D_{KL}(q_\varphi(z) || p(z))$$

Approches de démêlement (*disentanglement*)



Imposition d'une structure particulière à l'espace latent :

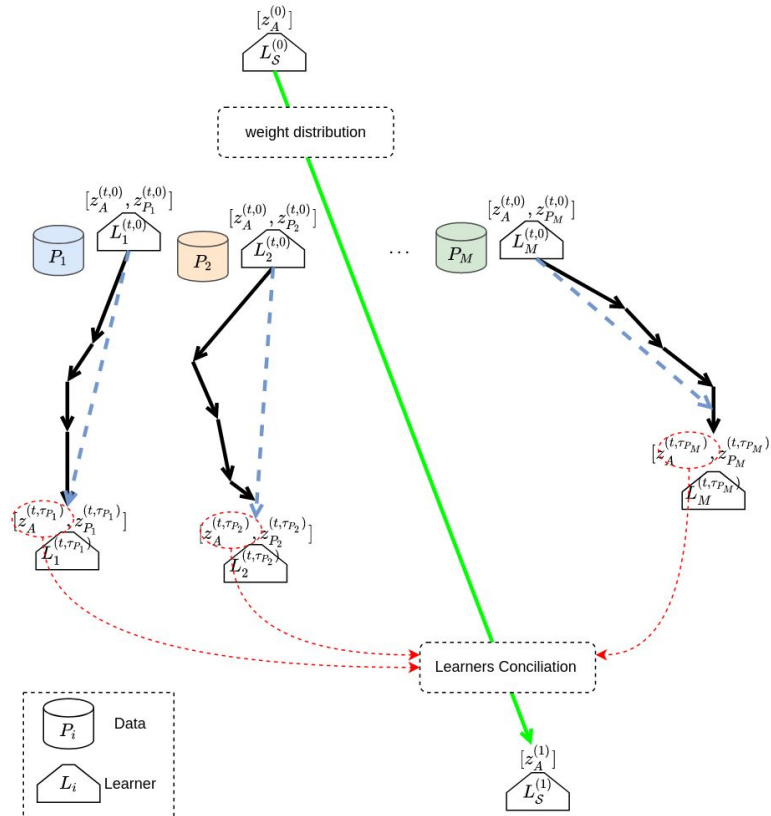
$$L(\theta, \varphi; x, z) = \mathbb{E}_{q_\varphi(z|x)} [\log p_\theta(x|z)] - \beta D_{KL}(q_\varphi(z|x) || p(z)) - \alpha D_{KL}(q_\varphi(z) || p(z))$$

Erreur de reconstruction

Divergence entre la distribution a posteriori et la structure cible

Contrainte imposant la parcimonie des représentations latentes

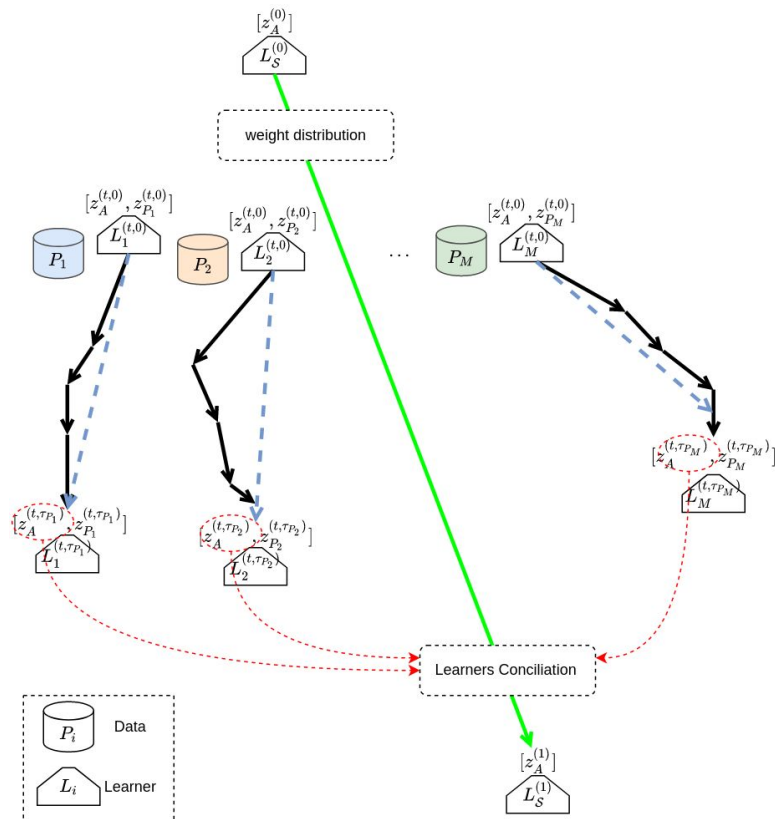
Apprentissage multi-niveaux



Apprenants locaux

- spécifiques aux positions du déploiement
- décomposition du signal/données en composantes spécifiques aux positions et universelles (mutualisables)

Apprentissage multi-niveaux



Apprenants locaux

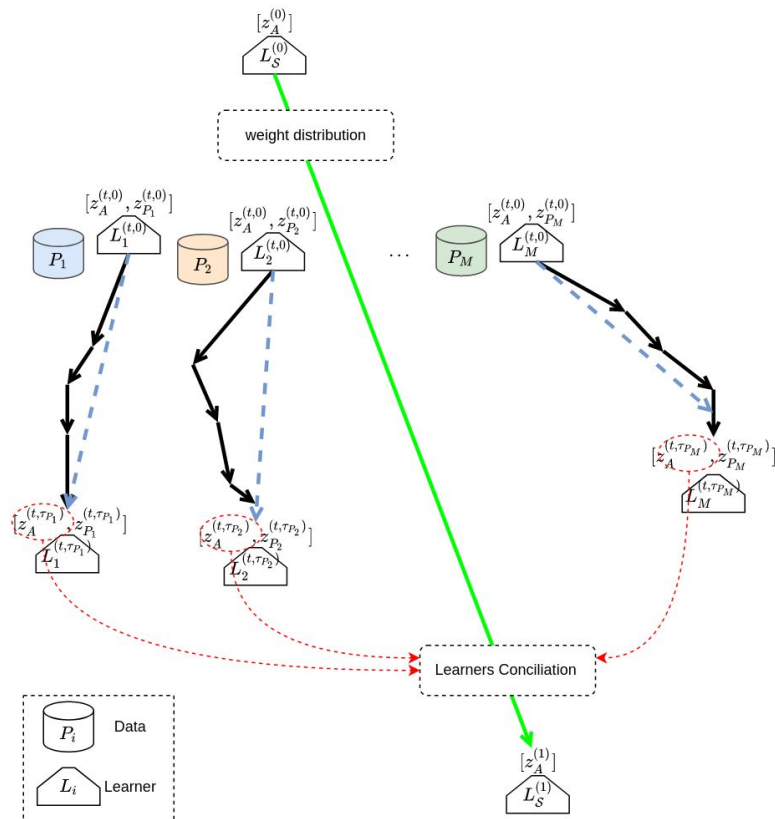
- spécifiques aux positions du déploiement
- décomposition du signal/données en composantes spécifiques aux positions et universelles (mutualisables)

L'objectif de l'apprenant local L_p peut être formalisé comme le risque empirique sur la distribution des données de la position p :

$$f_p(w_p) = \mathbb{E}_{\xi_p}[\tilde{f}_p(w_p; \xi_p)]$$

où ξ_p est un échantillon de données aléatoires tiré selon la distribution de la position p et $\tilde{f}_p(w_p; \xi_p)$ est une fonction objective correspondant à cet échantillon. w_p l'ensemble des poids de l'apprenant.

Apprentissage multi-niveaux



Apprenants locaux

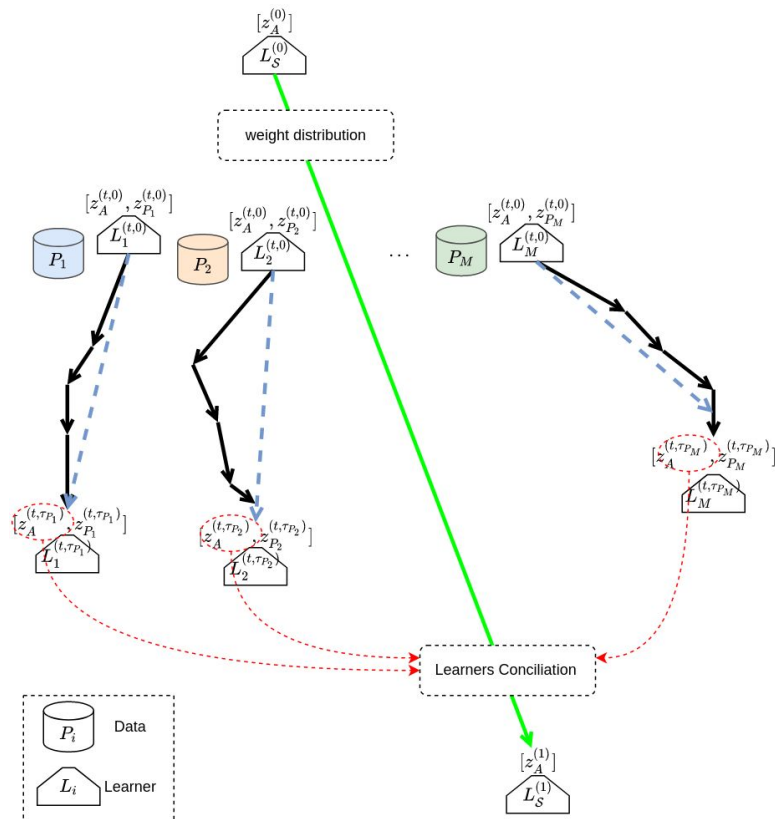
- spécifiques aux positions du déploiement
- décomposition du signal/données en composantes spécifiques aux positions et universelles (mutualisables)

Apprenant (central) référentiel

- conciliation des différentes perspectives

$$\min_{w \in \mathbb{R}^d} \left\{ F(w) := \sum_{p=1}^M \alpha_p \times f_p(w_p) \right\} \text{ with } \sum_{p=1}^M \alpha_p = 1$$

Apprentissage multi-niveaux



Apprenants locaux

- spécifiques aux positions du déploiement
- décomposition du signal/données en composantes spécifiques aux positions et universelles (mutualisables)

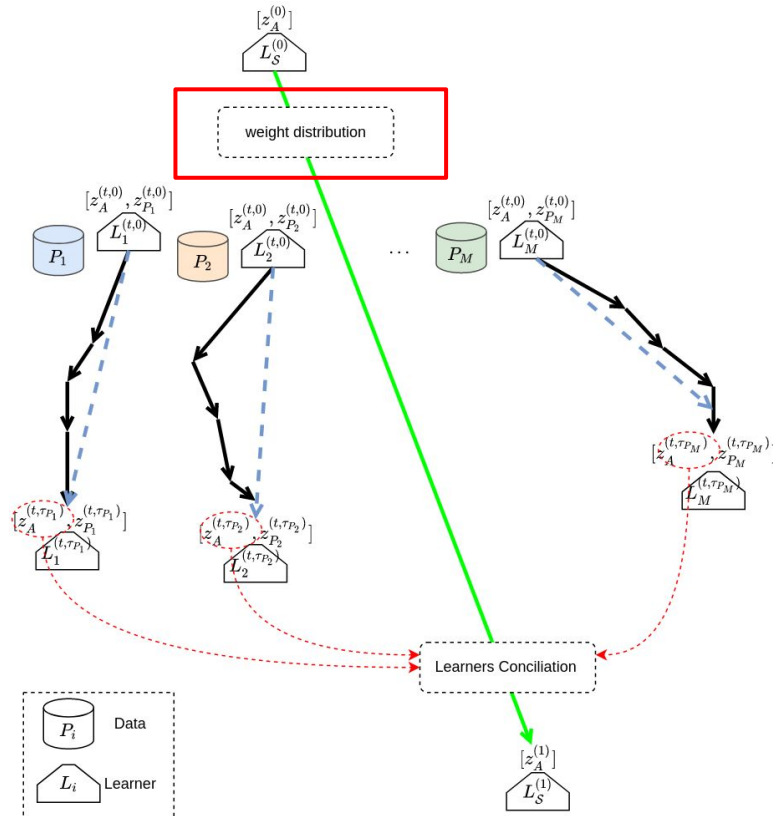
Apprenant (central) référentiel

- conciliation des différentes perspectives
- alignement des représentations universelles

$$\min_{w \in \mathbb{R}^d} \left\{ F(w) := \sum_{p=1}^M \alpha_p \times f_p(w_p) \right\} \text{ with } \sum_{p=1}^M \alpha_p = 1$$

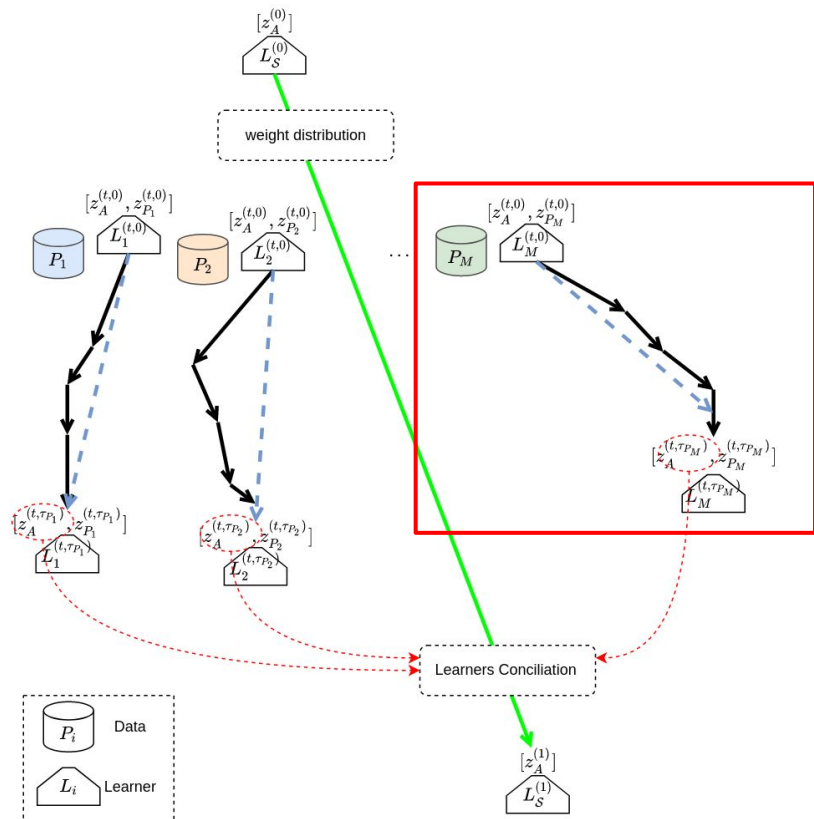
$$\min_{w \in \mathbb{R}^d} \left\{ F(w) = \frac{1}{M} \sum_{p=1}^M F_p(w_p) \right\}, F_p(w_p) = \min_{w \in \mathbb{R}^d} \left\{ f_p(w_p) + \lambda R(z_{iA}, z_A^{(t)}) \right\}$$

Abstraction de la Position des Capteurs



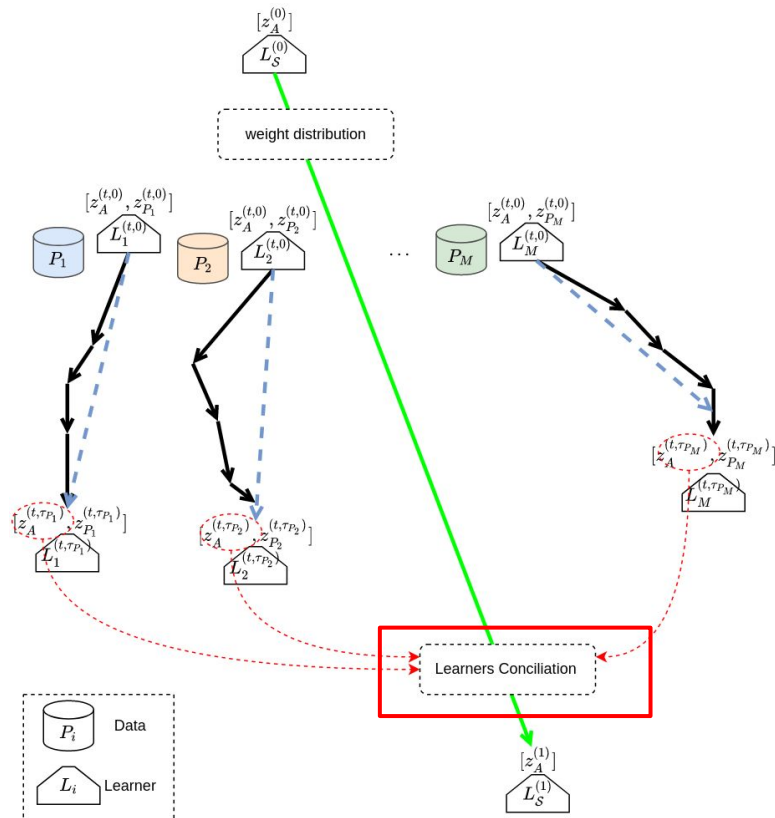
Initialisation des poids de l'apprenant référentiel et leur distribution aux apprenants locaux

Abstraction de la Position des Capteurs



A l'étape t du cycle de communication, chaque apprenant local exécute indépendamment τ_p itérations du solveur local basé sur la descente de gradient stochastique (SGD) à partir du modèle global courant $L^{(t,0)}$ jusqu'à l'étape $L^{(t,\tau_p)}$ afin d'optimiser sa propre fonction objective locale (flèches noires).

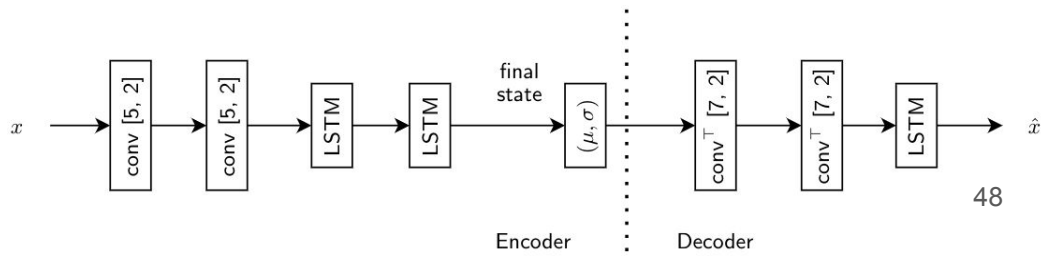
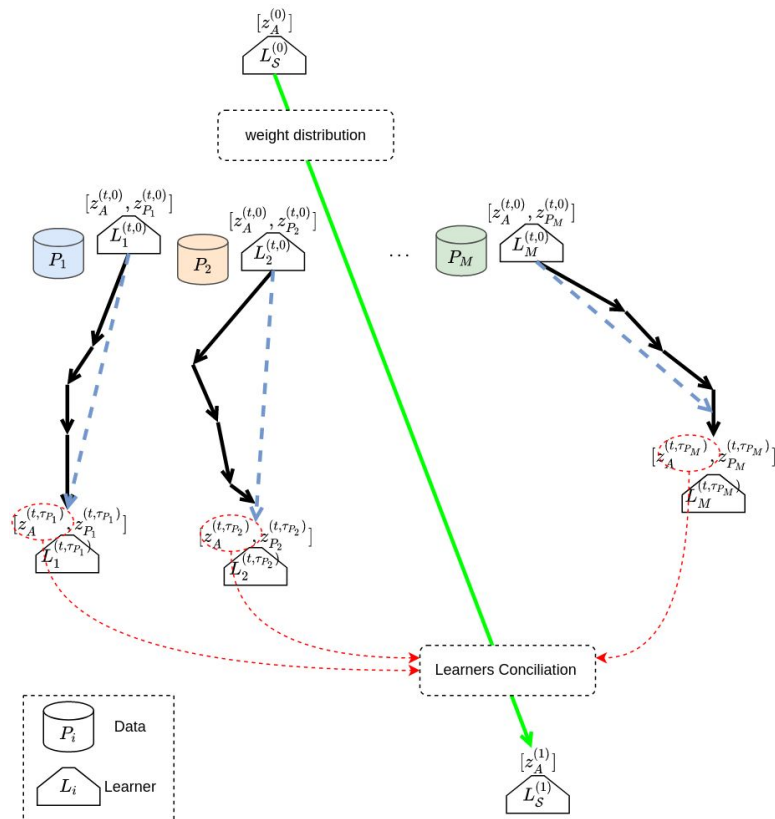
Abstraction de la Position des Capteurs



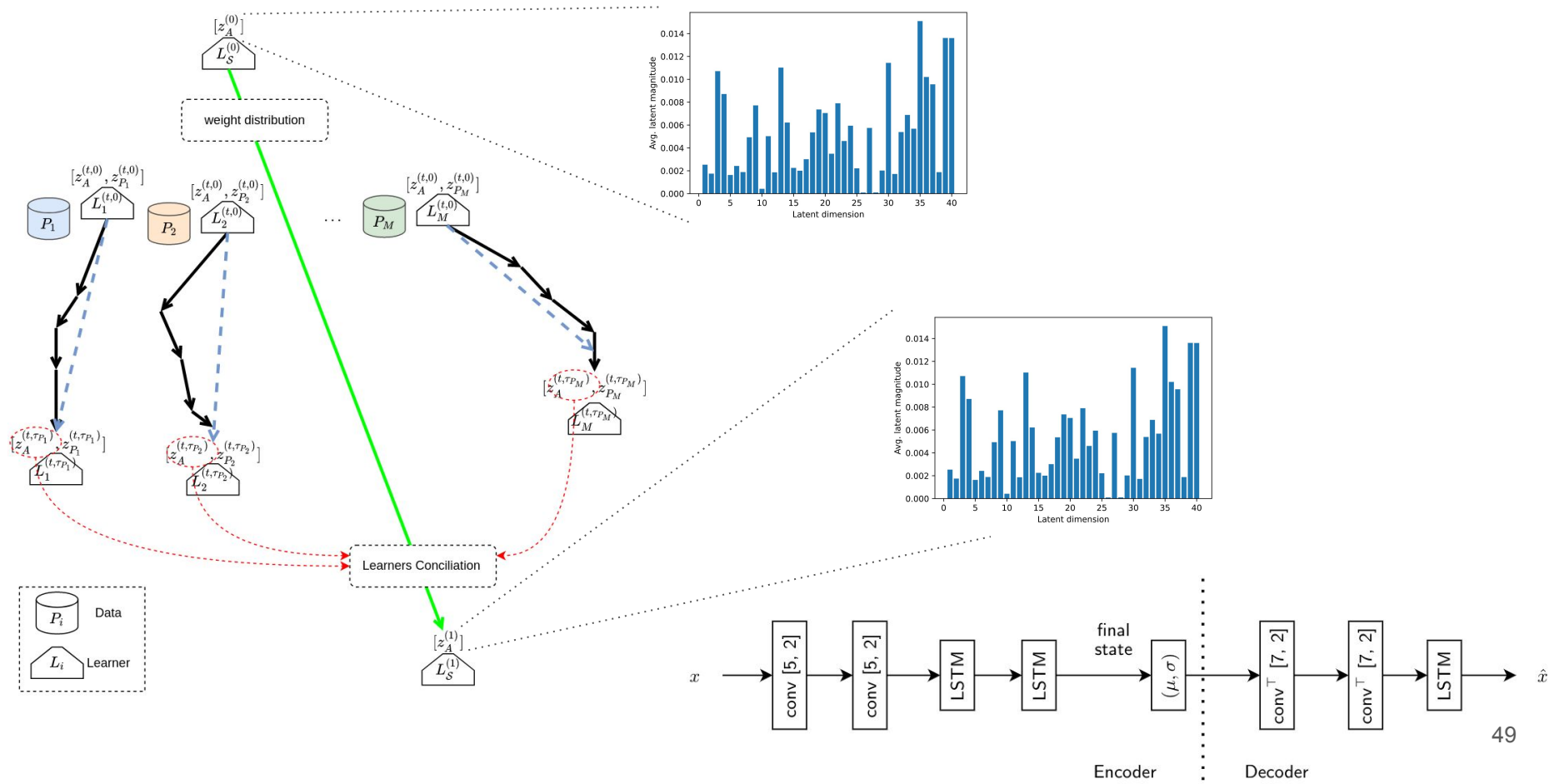
$$w^{(t+1,0)} - w^{(t,0)} = \sum_{p=1}^M \alpha_p \Delta_p^{(t)} = - \sum_{p=1}^M \alpha_p \cdot \eta \sum_{k=0}^{\tau_p-1} g_p(w_p^{(t,k)})$$

où $w_p^{(t,k)}$ désigne le modèle du client p après la k -ième mise à jour locale dans le t -ième cycle de conciliation et $\Delta_p^{(t)} = w_p^{(t,\tau_p)} - w_p^{(t,0)}$ désigne la progression locale réalisée par le client p au cycle t . De plus, η est le taux d'apprentissage du client et g_p représente le gradient stochastique sur un mini-lot d'échantillons B .

Quelques résultats : impact de la conciliation

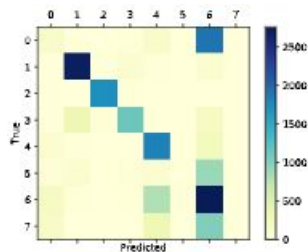


Quelques résultats : impact de la conciliation

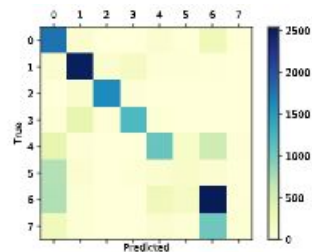


Quelques résultats : impact de la conciliation

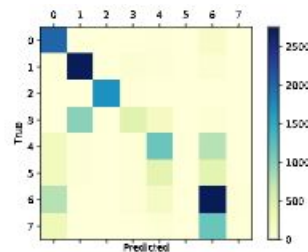
Config.	Recognition Performances \pm std.			
	<i>Bag</i>	<i>Hand</i>	<i>Hips</i>	<i>Torso</i>
Baseline (no sep.)	63.79 \pm .0089	63.86 \pm .0014	65.70 \pm .0126	60.61 \pm .0072
Universal comp.				
w/o conciliation	66.17 \pm .0224	65.26 \pm .0147	66.12 \pm .0035	62.47 \pm .013
w/ conciliation	66.97 \pm .016	67.8 \pm .0015	67.84 \pm .0354	63.12 \pm .01
Pos.-specific comp.				
w/o conciliation	64.2 \pm .3	66.17 \pm .007	67.9 \pm .0026	61.32 \pm .087
w/ conciliation	65.66 \pm .029	68.94 \pm .03	70.45 \pm .07	61.15 \pm .029



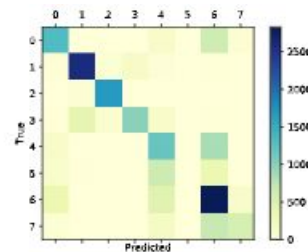
(a) $z_A; z_{Torso}$



(b) $z_A; z_{Hand}$



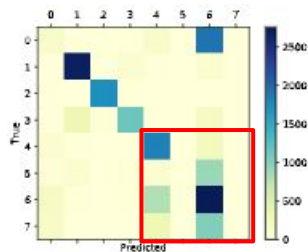
(c) $z_A; z_{Bag}$



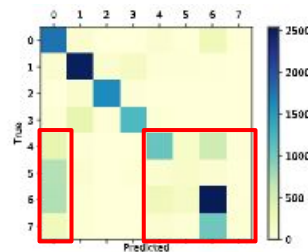
(d) $z_A; z_{Hips}$

Quelques résultats : impact de la conciliation

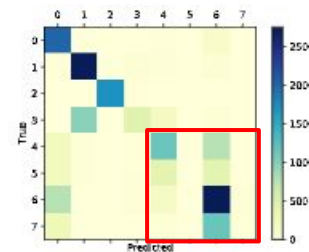
Config.	Recognition Performances \pm std.			
	<i>Bag</i>	<i>Hand</i>	<i>Hips</i>	<i>Torso</i>
Baseline (no sep.)	63.79 \pm .0089	63.86 \pm .0014	65.70 \pm .0126	60.61 \pm .0072
Universal comp.				
w/o conciliation	66.17 \pm .0224	65.26 \pm .0147	66.12 \pm .0035	62.47 \pm .013
w/ conciliation	66.97 \pm .016	67.8 \pm .0015	67.84 \pm .0354	63.12 \pm .01
Pos.-specific comp.				
w/o conciliation	64.2 \pm .3	66.17 \pm .007	67.9 \pm .0026	61.32 \pm .087
w/ conciliation	65.66 \pm .029	68.94 \pm .03	70.45 \pm .07	61.15 \pm .029



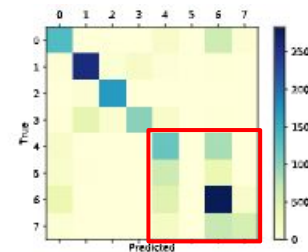
(a) $z_A; z_{Torso}$



(b) $z_A; z_{Hand}$



(c) $z_A; z_{Bag}$



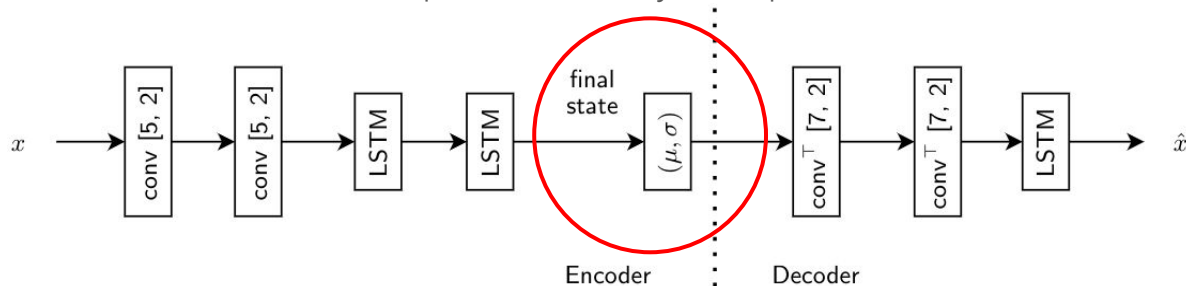
(d) $z_A; z_{Hips}$

Quelques résultats : configurations d'inférence

Config.	Best Config.	Recogn. Perf. \pm std.	mean \pm std.
Baselines			
Concat. fusion	-	-	60.24 \pm .014
Corr. Alignment	-	-	63.79 \pm .032
Activities			
<i>Still</i>	$z_{hi}; z_t$	85.77 \pm 0.016	83.26 \pm 0.7
<i>Walk</i>	$z_A; z_{ha}$	88.54 \pm 0.07	86.74 \pm 0.058
<i>Run</i>	z_{ha}	90.51 \pm 0.016	89.46 \pm 0.03
<i>Bike</i>	$z_A; z_{hi}$	85.62 \pm 0.2	83.22 \pm 0.086
<i>Car</i>	$z_A; z_{ha}$	78.24 \pm 0.058	77.14 \pm 0.2
<i>Bus</i>	z_{ha}	78.08 \pm 0.022	75.17 \pm 0.004
<i>Train</i>	$z_{hi}; z_{hi}$	76.13 \pm 0.175	74.88 \pm 0.08
<i>Subway</i>	$z_A; z_{ha}; z_t$	75.89 \pm 0.009	74.07 \pm 0.006

Axes d'amélioration

Modélisation et incorporation des dynamiques du mouvement

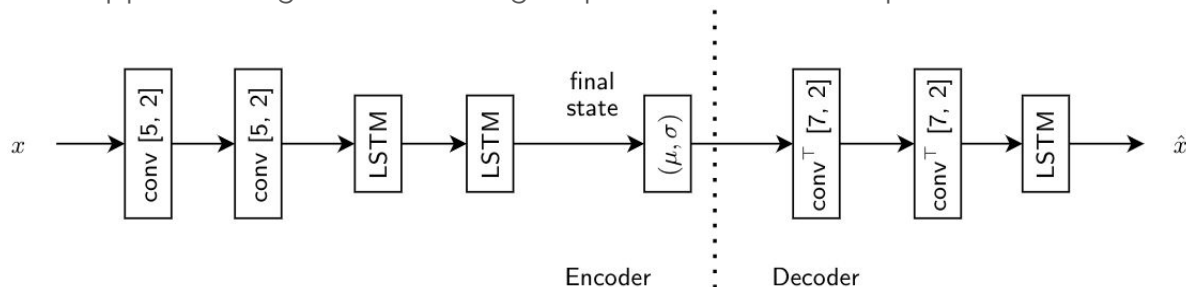


Un contrôle plus fin du processus d'apprentissage de représentations avec l'incorporation des primitives de mouvement pour chaque activité.

Les transitions entre les états du système dynamique sont explicitement encodées dans l'espace latent.

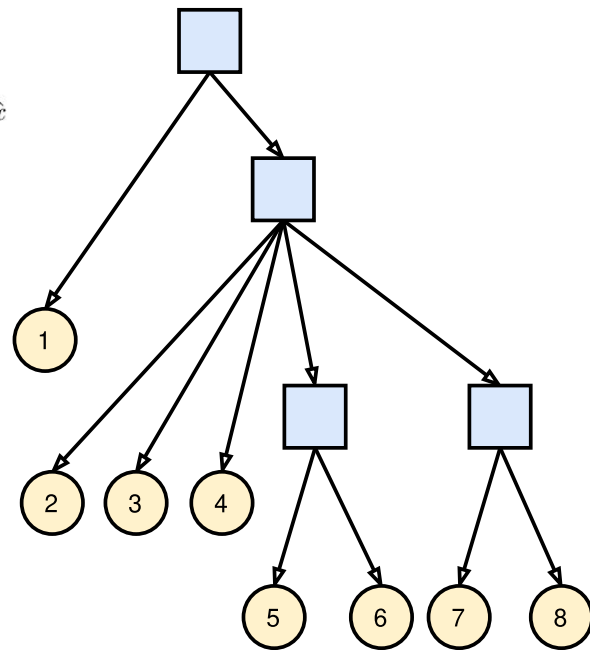
Axes d'amélioration

Apprentissage selon les regroupements de concepts



Un contrôle plus fin du processus d'apprentissage de représentations avec l'incorporation des primitives de mouvement pour chaque activité.

Les transitions entre les états du système dynamique sont explicitement encodées dans l'espace latent.



Massinissa Hamidi
LIPN-UMR CNRS 7030, Univ. Sorbonne Paris Nord

