Federated Multi-task Learning

知的学習論研究室　M1 206t244t 中島弘樹

目次

1,研究背景

2,前提知識

2-1, Federation Learning

2-2, Multi-Task Learning

3, 提案手法 ~MOCHAについて~

4, Result

目次

1,研究背景

2,前提知識

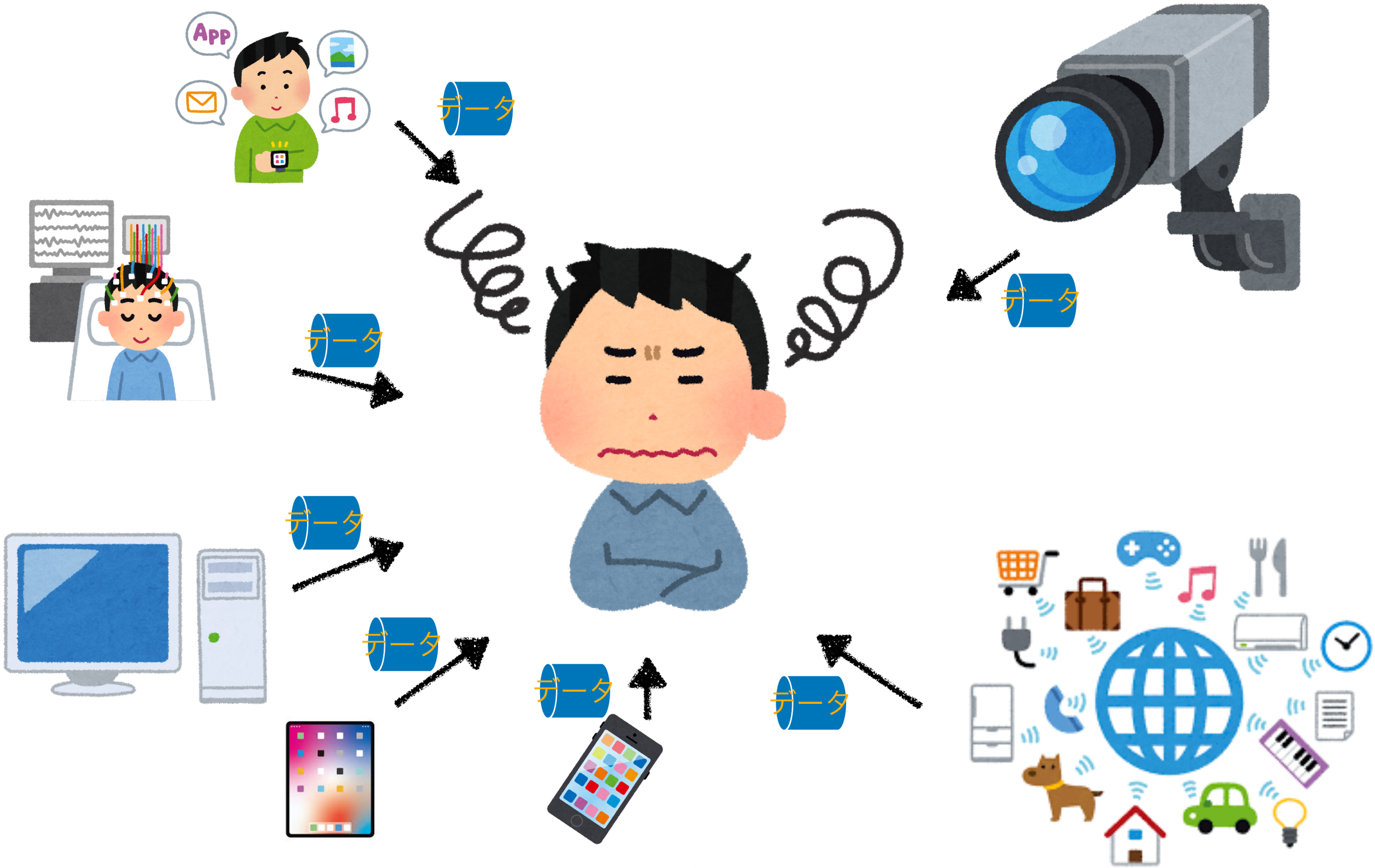
2-1, Federation Learning

2-2, Multi-Task Learning

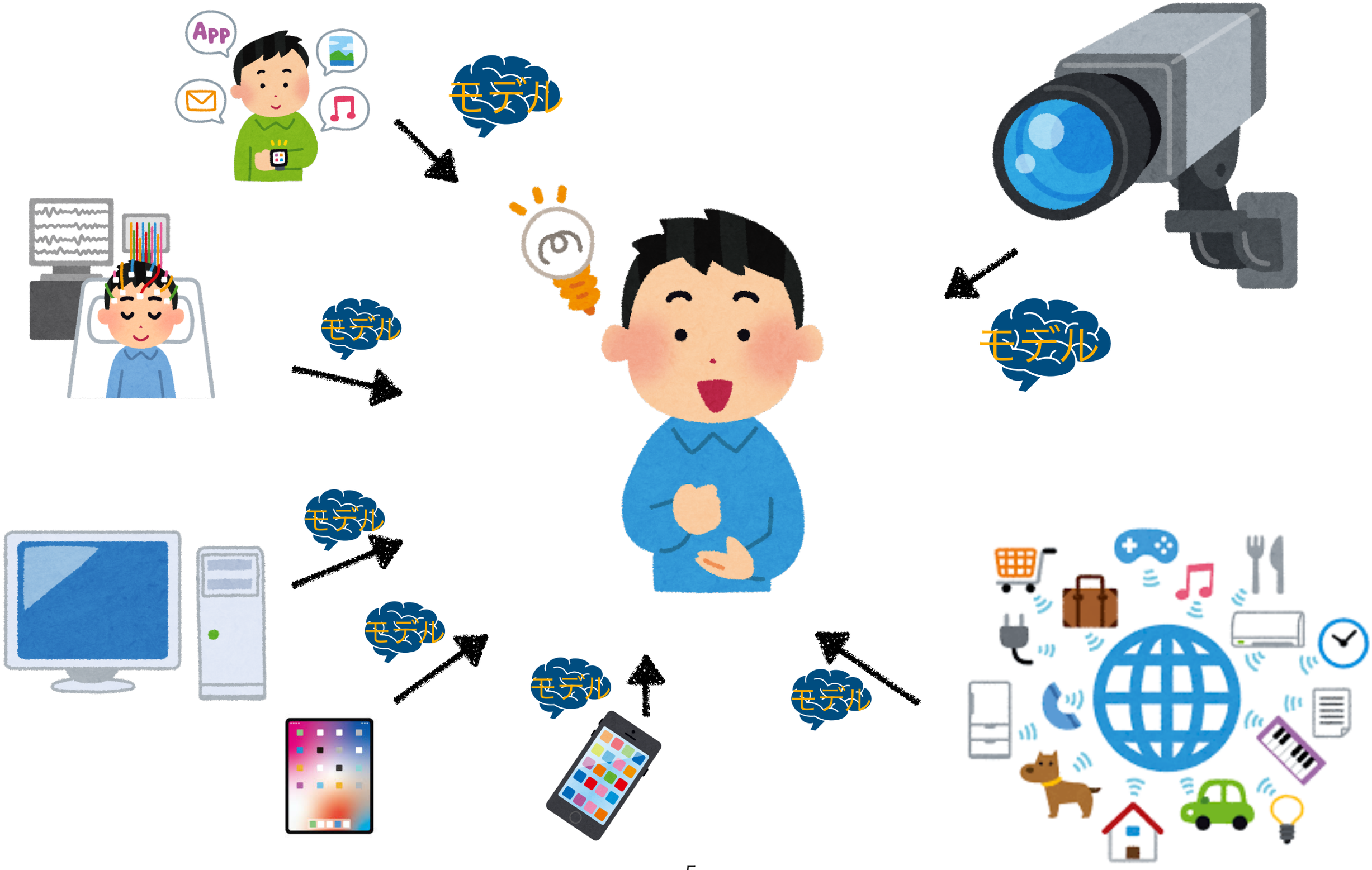
3, 提案手法 ~MOCHAについて~

4, Result

研究背景



研究背景



概要

1 : Federated Learning の二つの問題

1-1 :統計的課題 : データの不均一性に関する課題
1-2 :システム課題 : 通信やストラグラーに関する問題

２ : 統計的課題に対して Multi-Task Learningが有効！！

３ : システム課題に対して最適化手法 “MOCHA” の提案！！

目次

1,研究背景

2,前提知識

2-1, Federation Learning

2-2, Multi-Task Learning

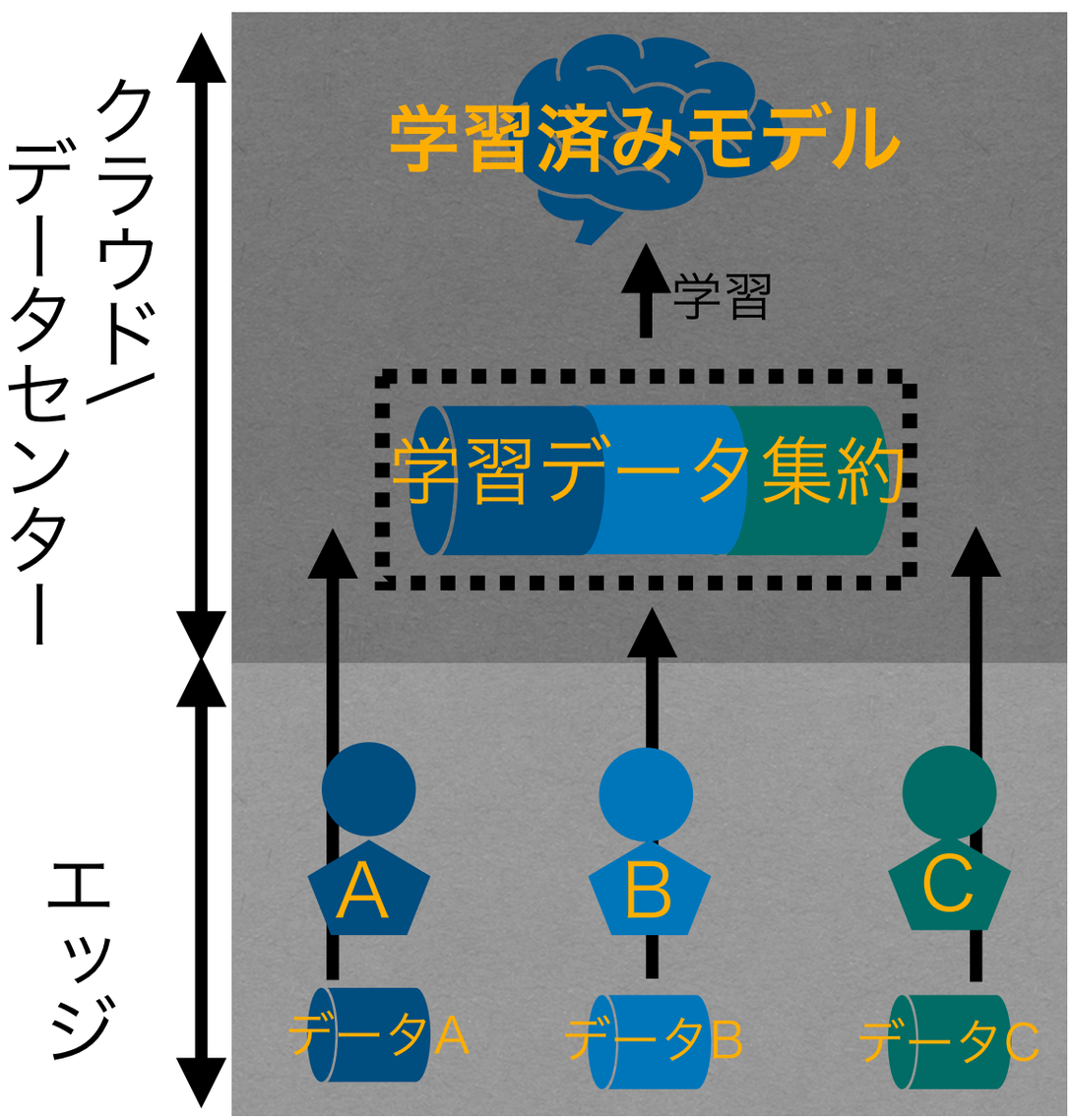
3, 提案手法 ~MOCHAについて~

4, Result

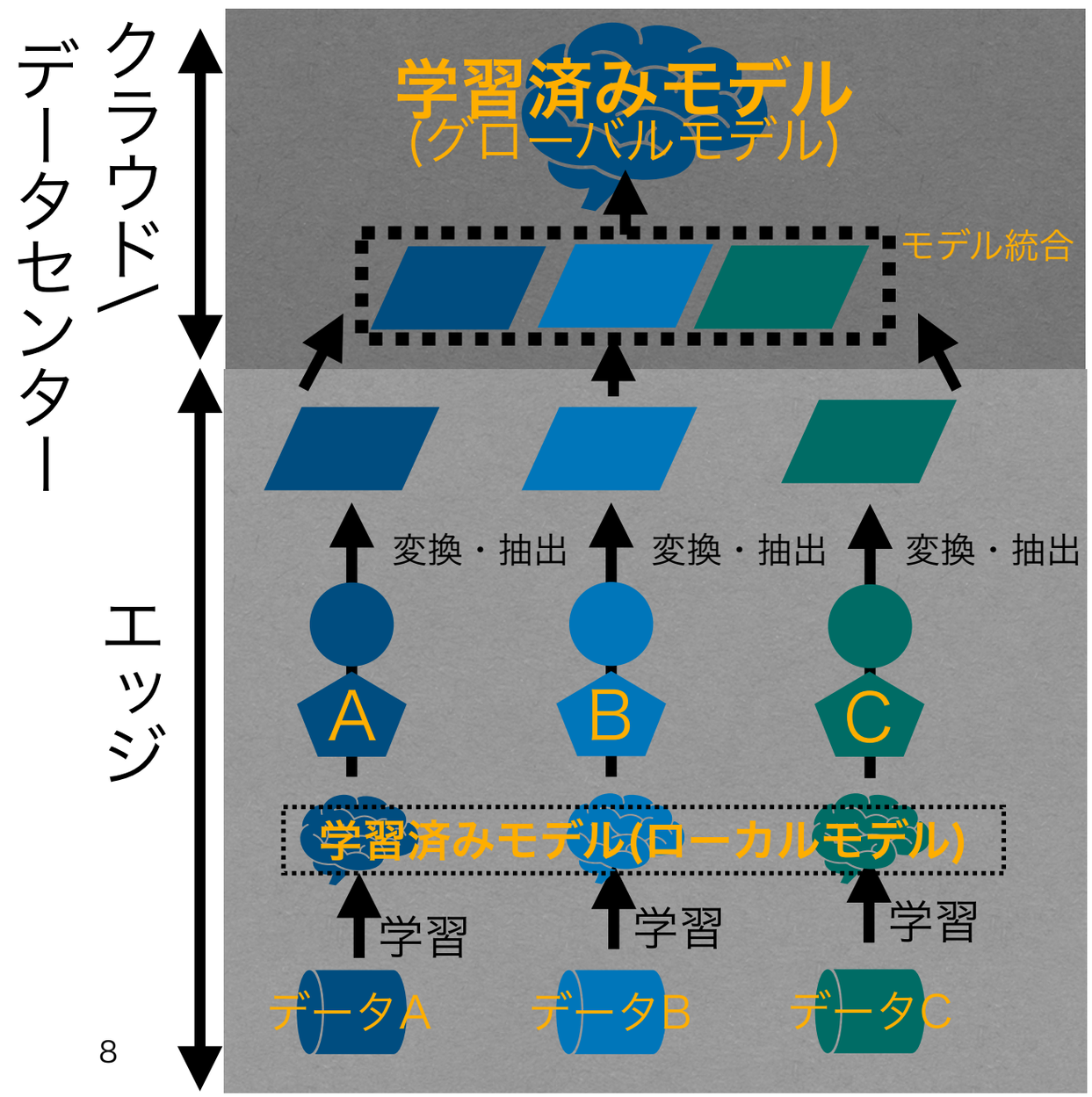
Federation Learning(FL) とは

：エッジ端末が持つデータを端末に保持したまま
利用する分散学習方式の1つ

一般的な機械学習のトレーニング

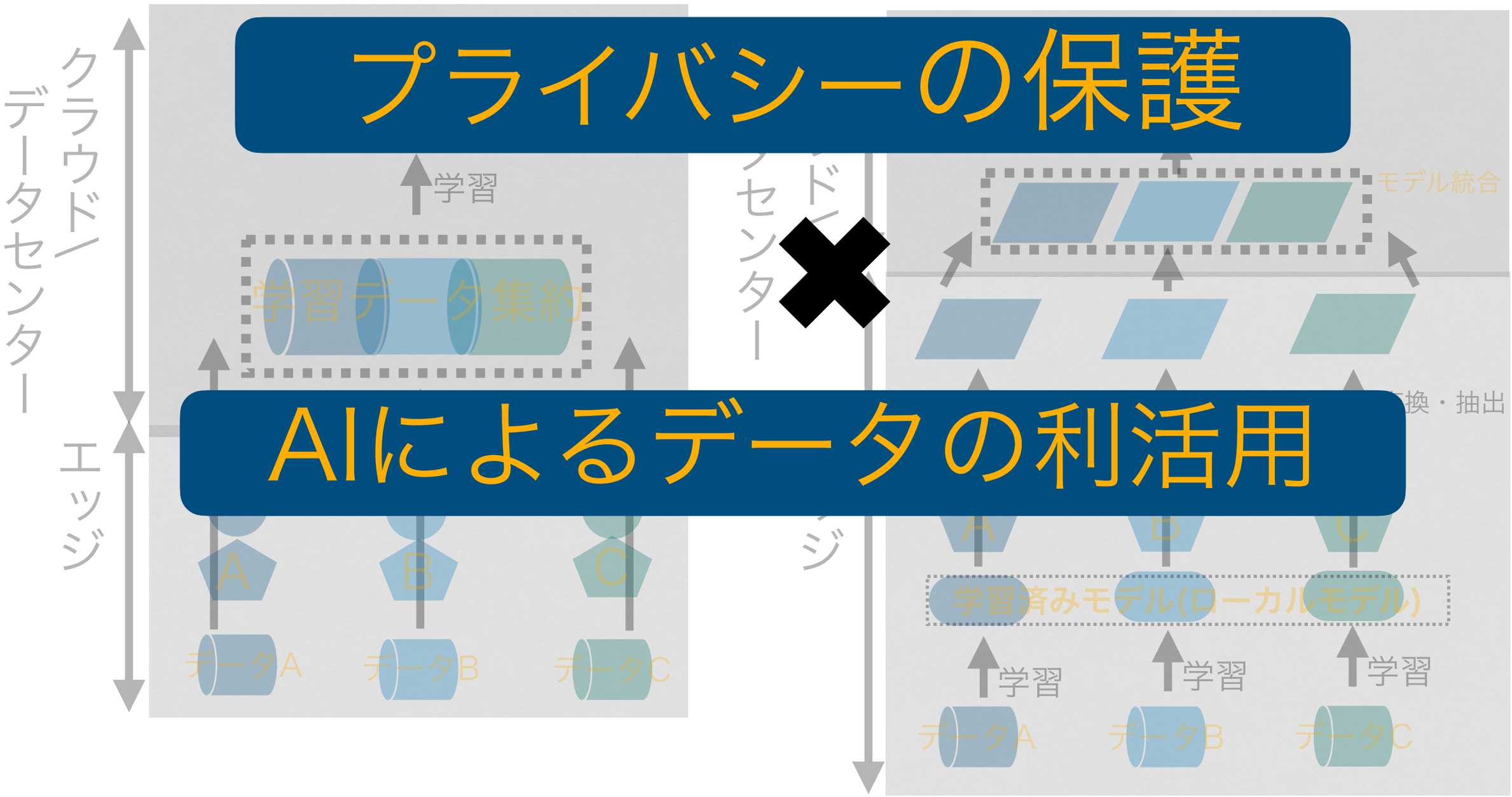


Federation Learningのトレーニング



Federation Learning とは

一般的な機械学習のトレーニング　Federation Learningのトレーニング



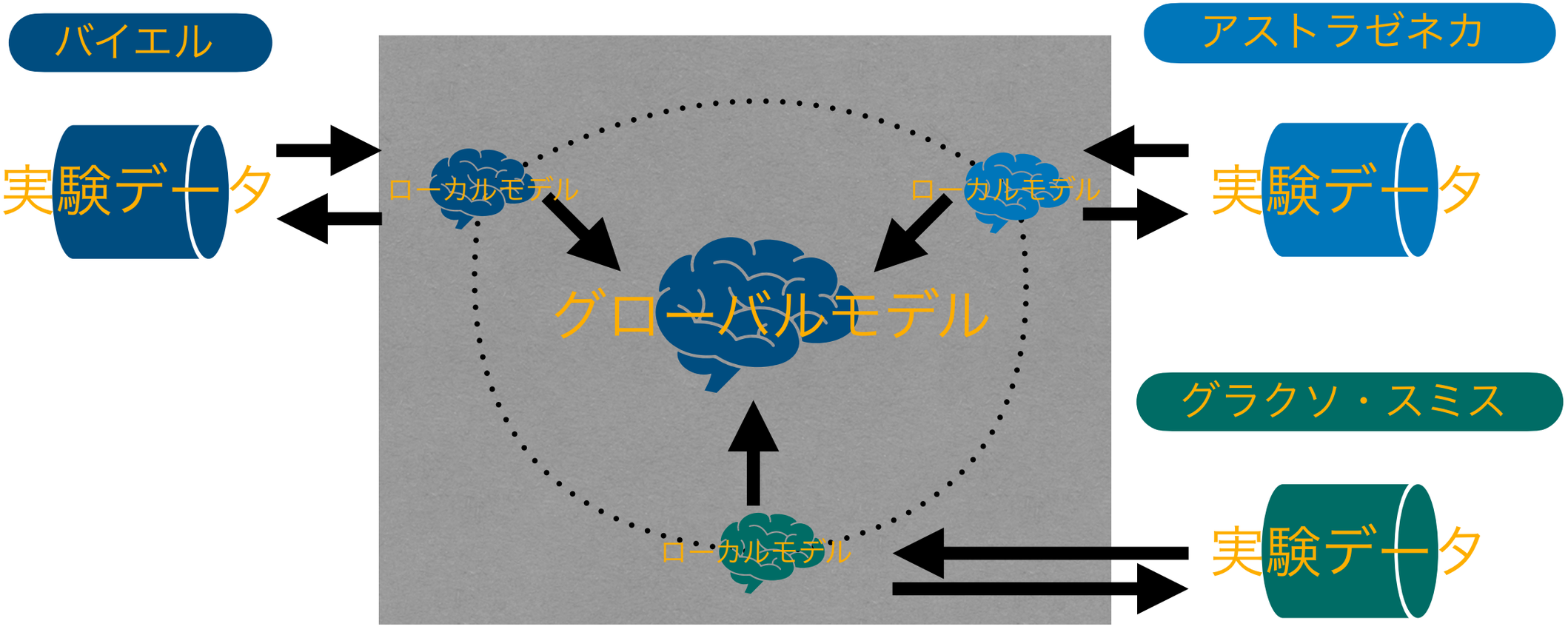
プライバシーの保護

AIによるデータの利活用

Federation Learning 例①

MELLODDY：創薬プロセスのイノベーション
(Machine Learning Ledger Orchestration for Drug Discovery)

：Federation Learningとブロックチェーンの技術を融合し、
膨大なデータの相互利用による創薬の開発研究を行うプロジェクト



Federation Learning 問題点

1：統計的課題：データの不均一性に関する課題

2：システム課題：通信やストラグラーに関する課題

ストラグラー: エッジ端末の性能差による計算時間の差

目次

1,研究背景

2,前提知識

2-1, Federation Learning

2-2, Multi-Task Learning

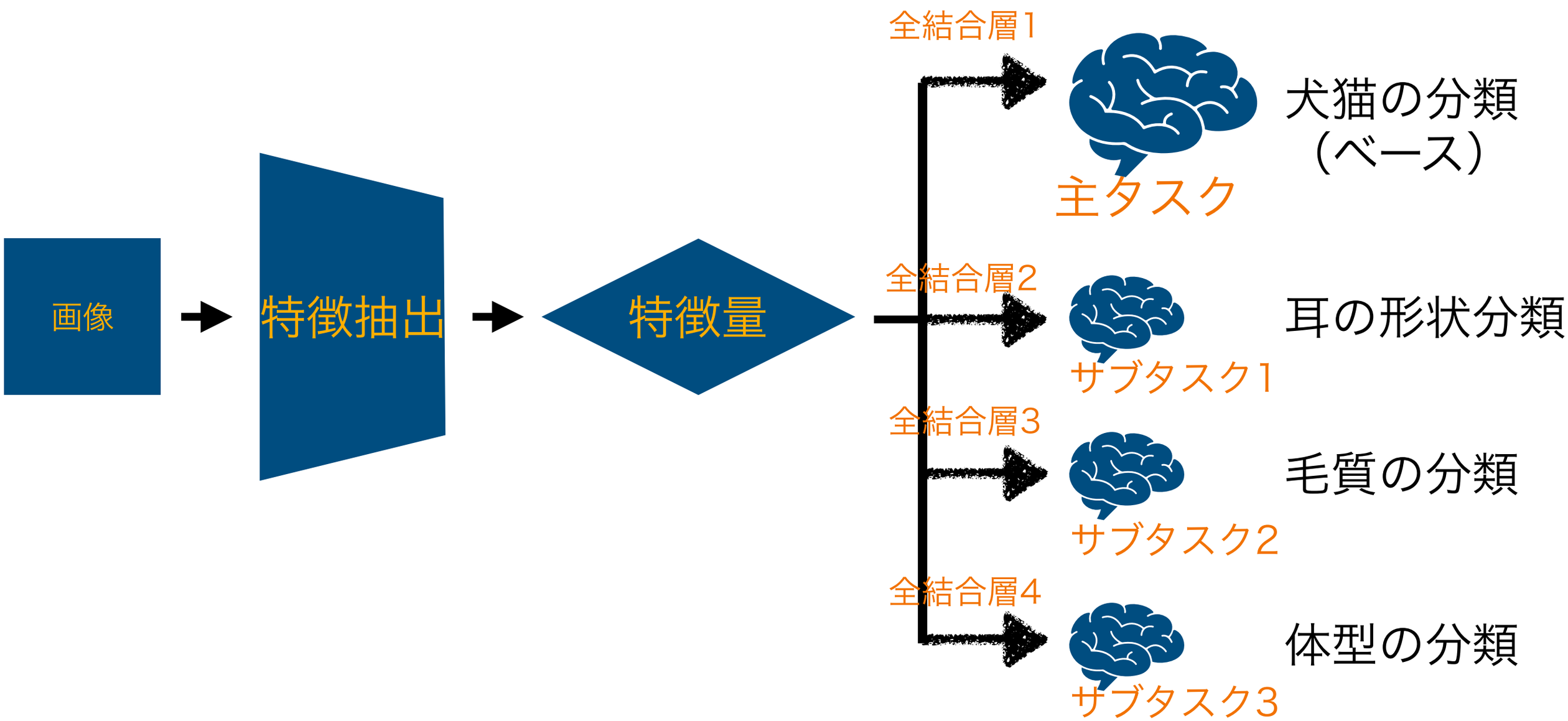
3, 提案手法 ~MOCHAについて~

4, Result

Multi-Task Learning

Multi-Task Learning：

一つのモデルで複数のタスクを解くことでモデルの汎化性能を向上



Multi-Task Learning

一般的なMulti-Task Learning

$$min(\sum_{t=1}^{m}\sum_{i=1}^{n_t} l_t(w_t^T x_t^i, y_t^i) + R(W,\Omega)) \quad \cdot\cdot\cdot(1)$$

$$W = [w_1, w_2, ..., w_m]$$

m : タスク数　n : データ数　**W,Ω : 重み**

L : 何かしらの損失関数(SVMとか)

R : 正則化　λ: 正則化パラメータ

W : データに対する重み　Ω : タスクに対する重み

目次

1,研究背景

2,前提知識

2-1, Federation Learning

2-2, Multi-Task Learning

3, 提案手法 ~MOCHAについて~

4, Result

“MOCHA”

Federated Learning : 各エッジ(ノード)にデータ X が存在

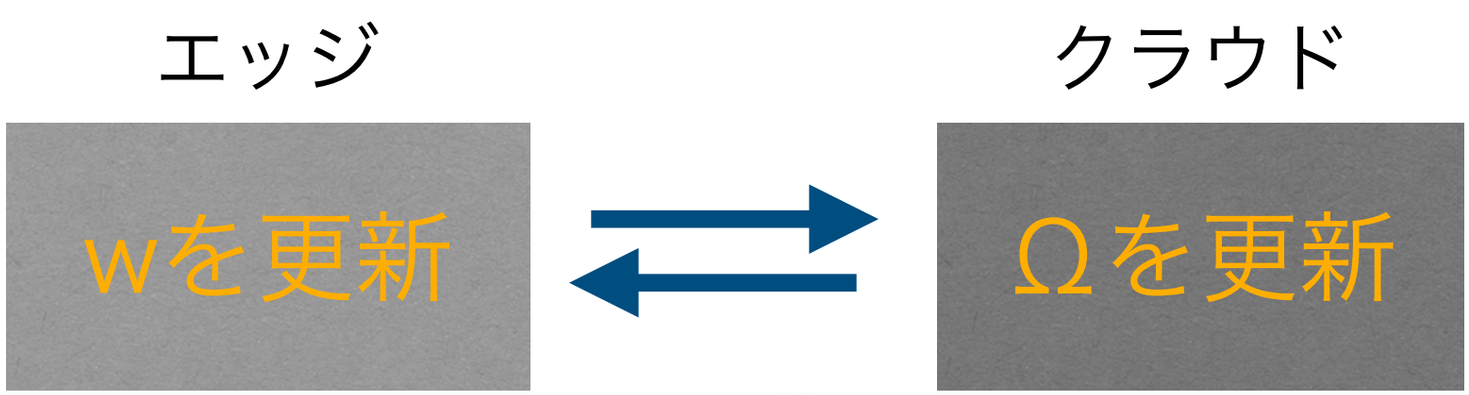
$$X = [x_1, x_2, ..., x_m]$$

$$min(\sum_{t=1}^{m}\sum_{i=1}^{n_t} l_t(w_t^T x_t^i, y_t^i) + R(W, \Omega)) \quad \cdot\cdot\cdot (1)$$

→　重み　WとΩ を学習したい！！！

→　同時に学習は不可能！！！

→　相互的に学習しよう!!(Wに焦点を当てます)



“MOCHA” ～Wの更新～

・“W” の更新：グローバルな損失関数

・(1)式の “Ω” を固定し、

$$min(D(\alpha) = \sum_{t=1}^{m}\sum_{i=1}^{n_t} l_t(-\alpha_t^i) + R(X\alpha)) \quad ・・・(3)$$

・Federated Multi-Task Learning 問題点

→ノード(タスク)が分散していること

→各タスクで Subproblem を定義

Subproblem

・Subproblem の定義：アップデート情報保持

$$min(g_t^{\sigma}(\Delta\alpha_t; v_t, \alpha_t)$$

$$= \sum_{i=1}^{n_t} l_t(-\alpha_t^i - \Delta\alpha_t^i) + < w_t(\alpha), X_t\Delta\alpha_t > + \frac{\sigma^{'}}{2}||X_t\Delta\alpha_t||_{M_t}^2 + c(\alpha)) \quad \cdot\cdot\cdot (4)$$

$$c(\alpha) = \frac{1}{m}R(X\alpha) \qquad w_t(\alpha) = \nabla R(X\alpha)$$

$v = X\alpha$ ：vt はアップデート情報

通信やストラグラーへの対処

$$\theta_t^h := \frac{\mathcal{G}_t^{\sigma'}(\Delta\boldsymbol{\alpha}_t^{(h)}; \mathbf{v}^{(h)}, \boldsymbol{\alpha}_t^{(h)}) - \mathcal{G}_t^{\sigma'}(\Delta\boldsymbol{\alpha}_t^{\star}; \mathbf{v}^{(h)}, \boldsymbol{\alpha}_t^{(h)})}{\mathcal{G}_t^{\sigma'}(\mathbf{0}; \mathbf{v}^{(h)}, \boldsymbol{\alpha}_t^{(h)}) - \mathcal{G}_t^{\sigma'}(\Delta\boldsymbol{\alpha}_t^{\star}; \mathbf{v}^{(h)}, \boldsymbol{\alpha}_t^{(h)})}, \quad \cdot\cdot\cdot (5)$$

ストラグラー : controller θ

→ θ (0~1の値を取る)を定義することで対処

$\theta_t^h = 0$　g(・)の解が見つかり、更新できる状態

$\theta_t^h = 1$　イテレーションの間に進捗が生まれていない状態

Ex) エッジのバッテリー切れ、通信制限など

目次

1,研究背景

2,前提知識

2-1, Federation Learning

2-2, Multi-Task Learning

3, 提案手法 ~MOCHAについて~

4, Result

Dataset

Dataset

- Google Glass(GLEAM)

  →被験者38人のGoogle Glassから得られるデータ

- Human Activity Recognition

  →30人のスマートフォンから得られるデータ

- Vehicle Sensor

  →23個の車両から得られるデータ

性能評価

Model	Human Activity	Google Glass	Vehicle Sensor
Global	2.23 (0.30)	5.34 (0.26)	13.4 (0.26)
Local	1.34 (0.21)	4.92 (0.26)	7.81 (0.13)
MTL	**0.46 (0.11)**	**2.02 (0.15)**	**6.59 (0.21)**

表1: 各手法、各データセットにおける平均誤差

ストラグラーと通信

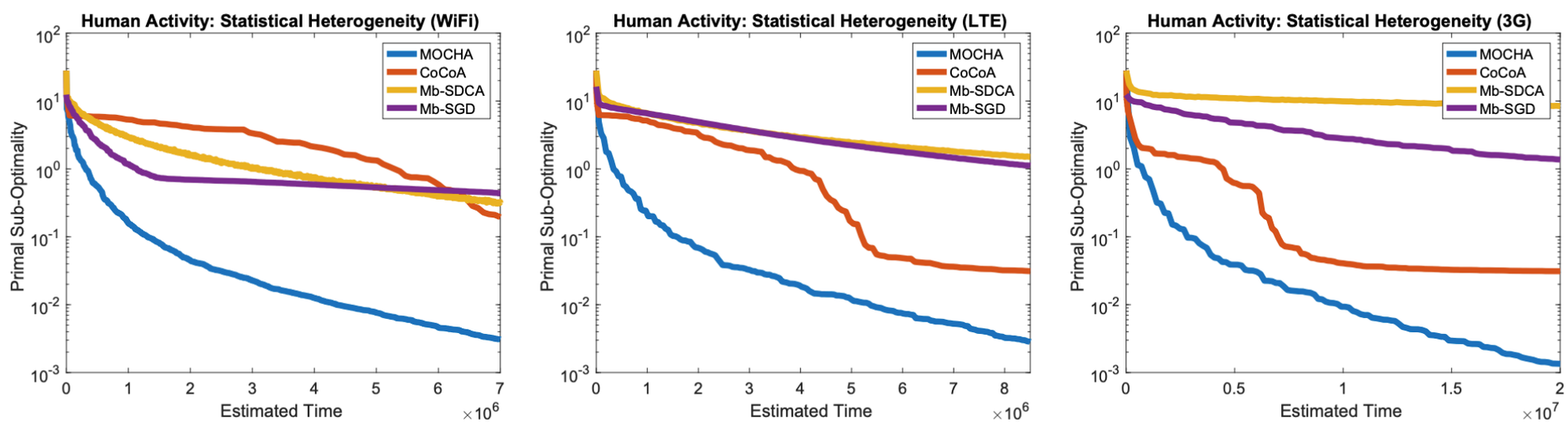


図1:　wの更新における通信環境の変化における各手法との比較

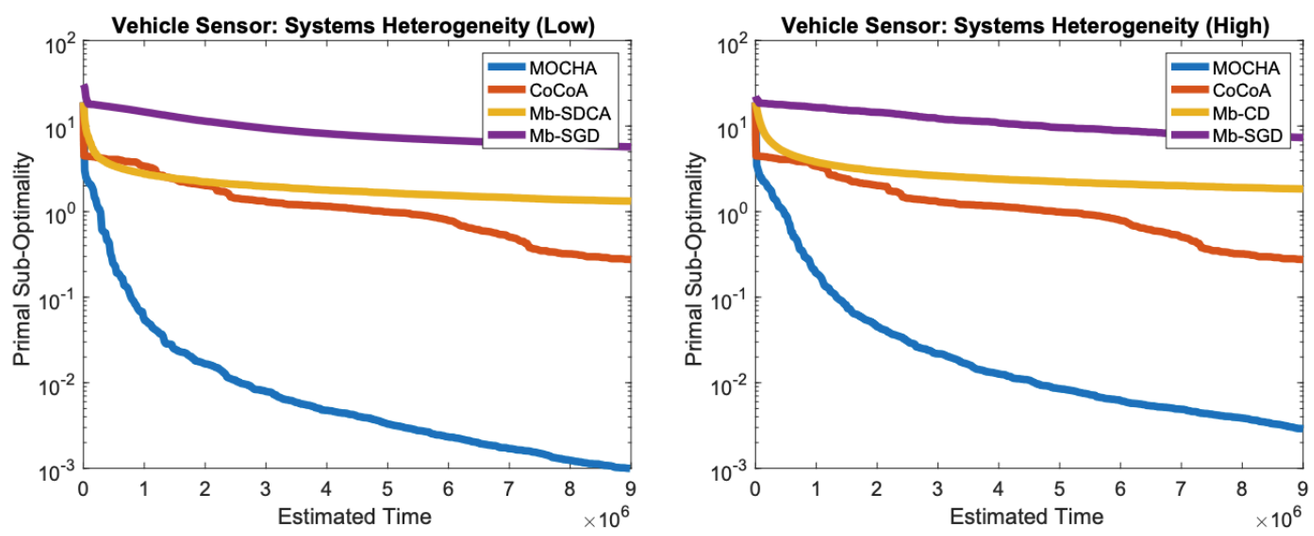


図2:　変動環境における各手法の比較

まとめ

1 : Federated Learning の二つの問題

1-1 :統計的課題 : データの不均一性に関する課題
1-2 :システム課題 : 通信やストラグラーに関する問題

2 : 統計的課題に対して Multi-Task Learningが有効！！

3 : システム課題に対して最適化手法 “MOCHA” の提案！！