UnB

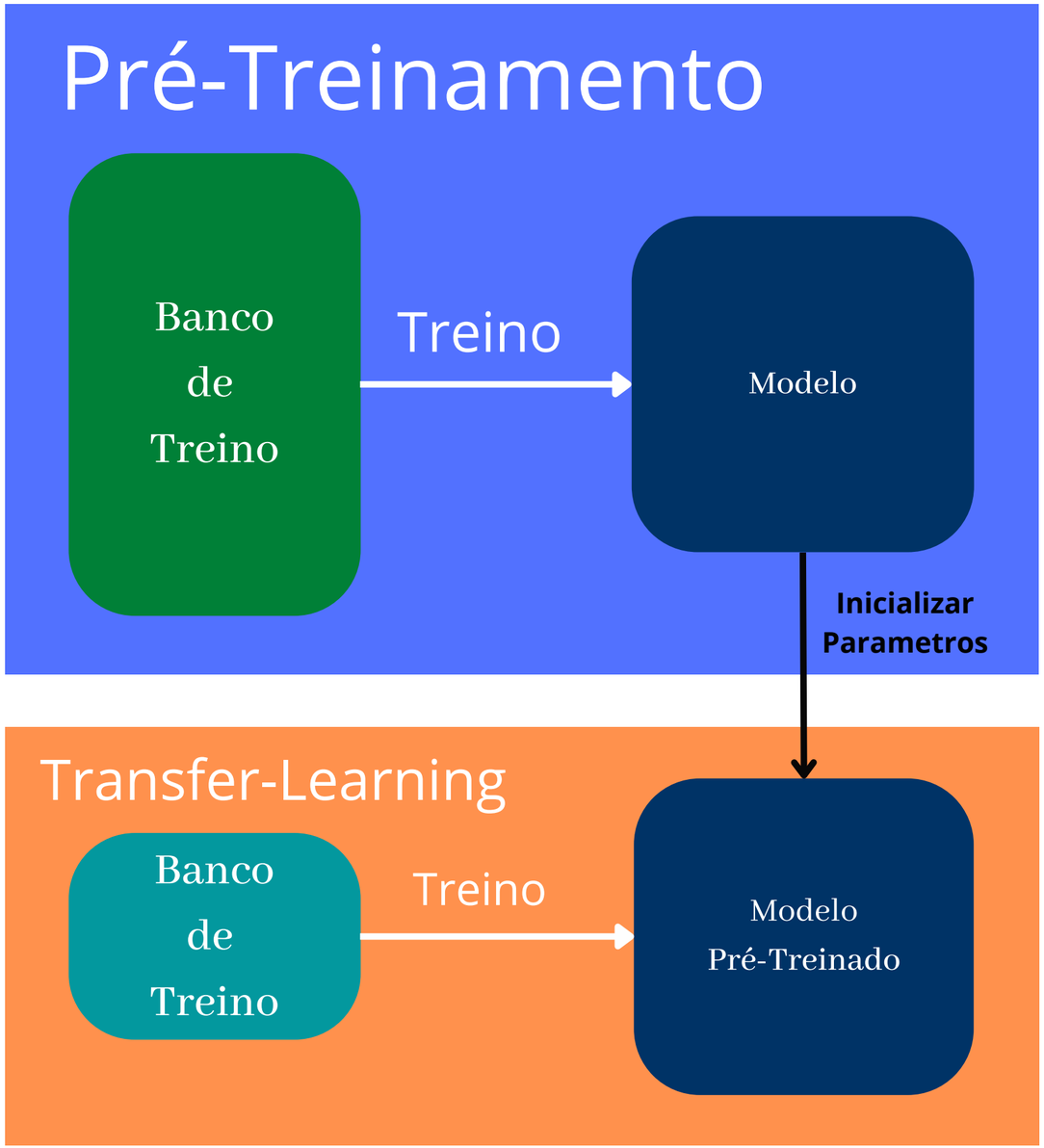
LAMFO

# Meta-Learning
# Few-Shot Learning

*Allan Faria*

08.2022

# Transfer-leaning



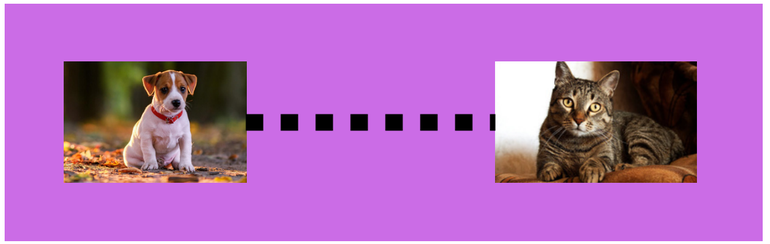

## Treino

$$\theta' = \min_{\theta} E_x[L_\theta]$$

# Tarefas

## 2-ways

**Tarefa 1: T1**



Cachorro: 1

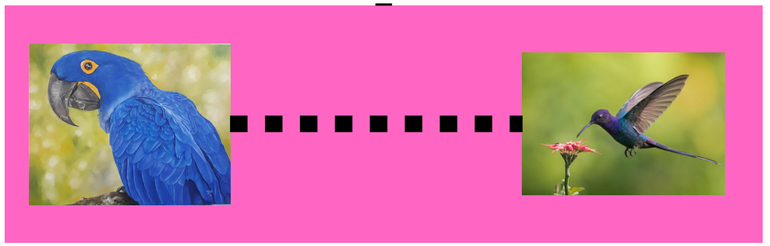
Gato: 0

**Tarefa 2: T2**

Cobra: 1

Largato: 0

**Tarefa n: Tn**



Arara: 1

Beija-Flor: 0

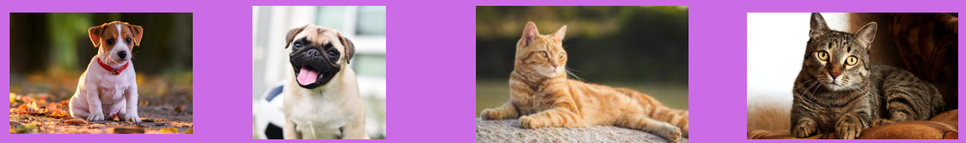
# Few-shot tasks

## 2-way-2-shots

## 2-shots

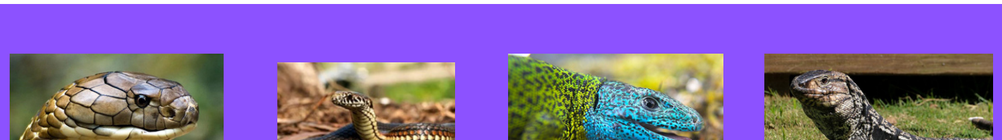
## 2-ways

**T1:**



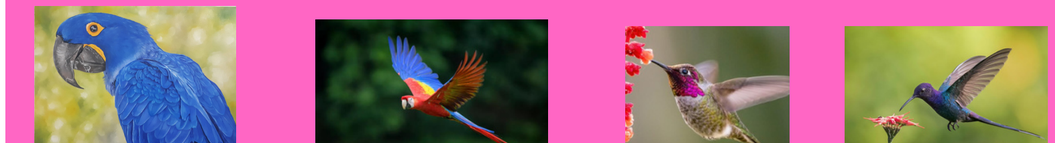
Cachorro: 1
Gato: 0

**T2:**



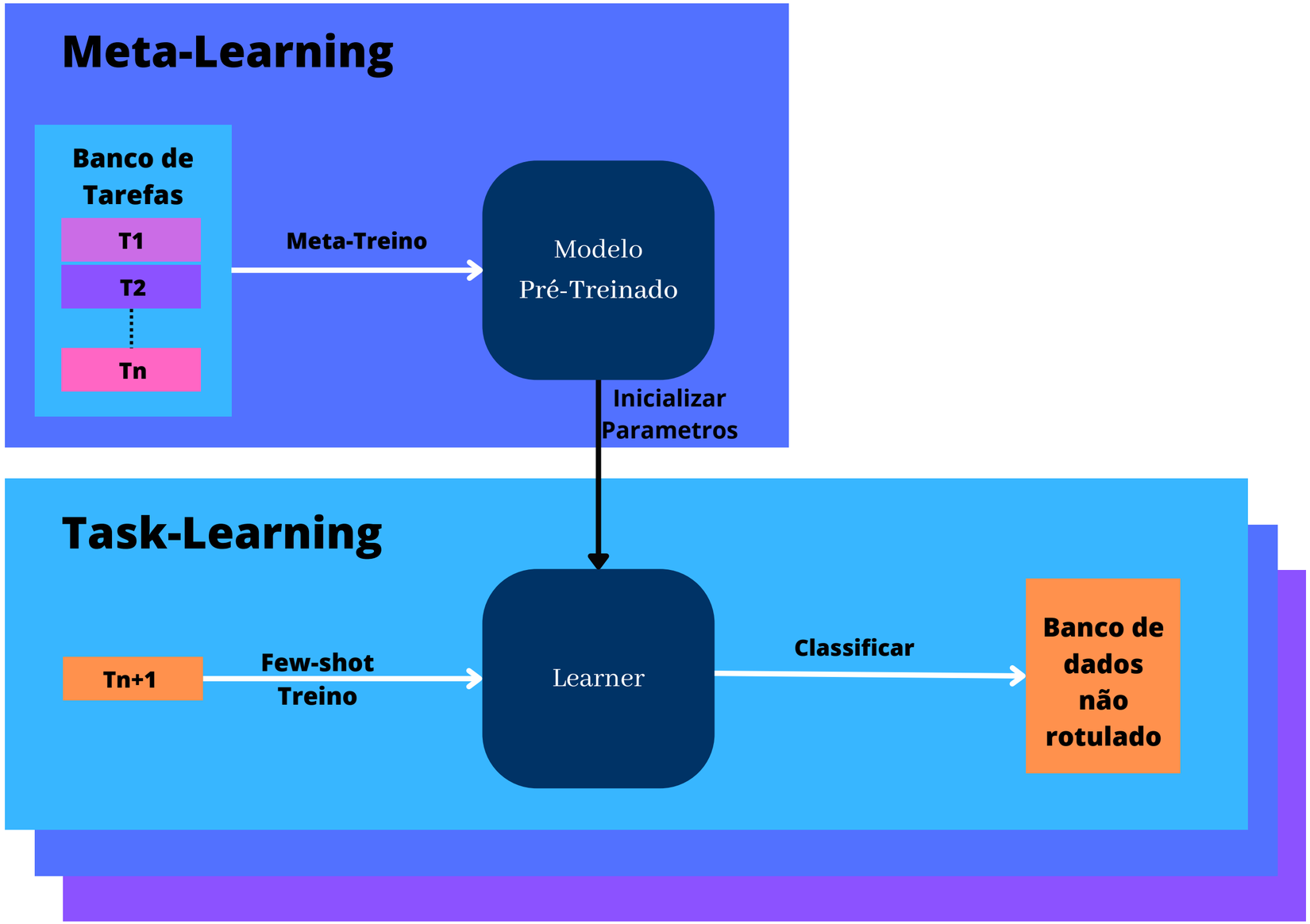
Cobra: 1
Largato: 0

**Tn:**

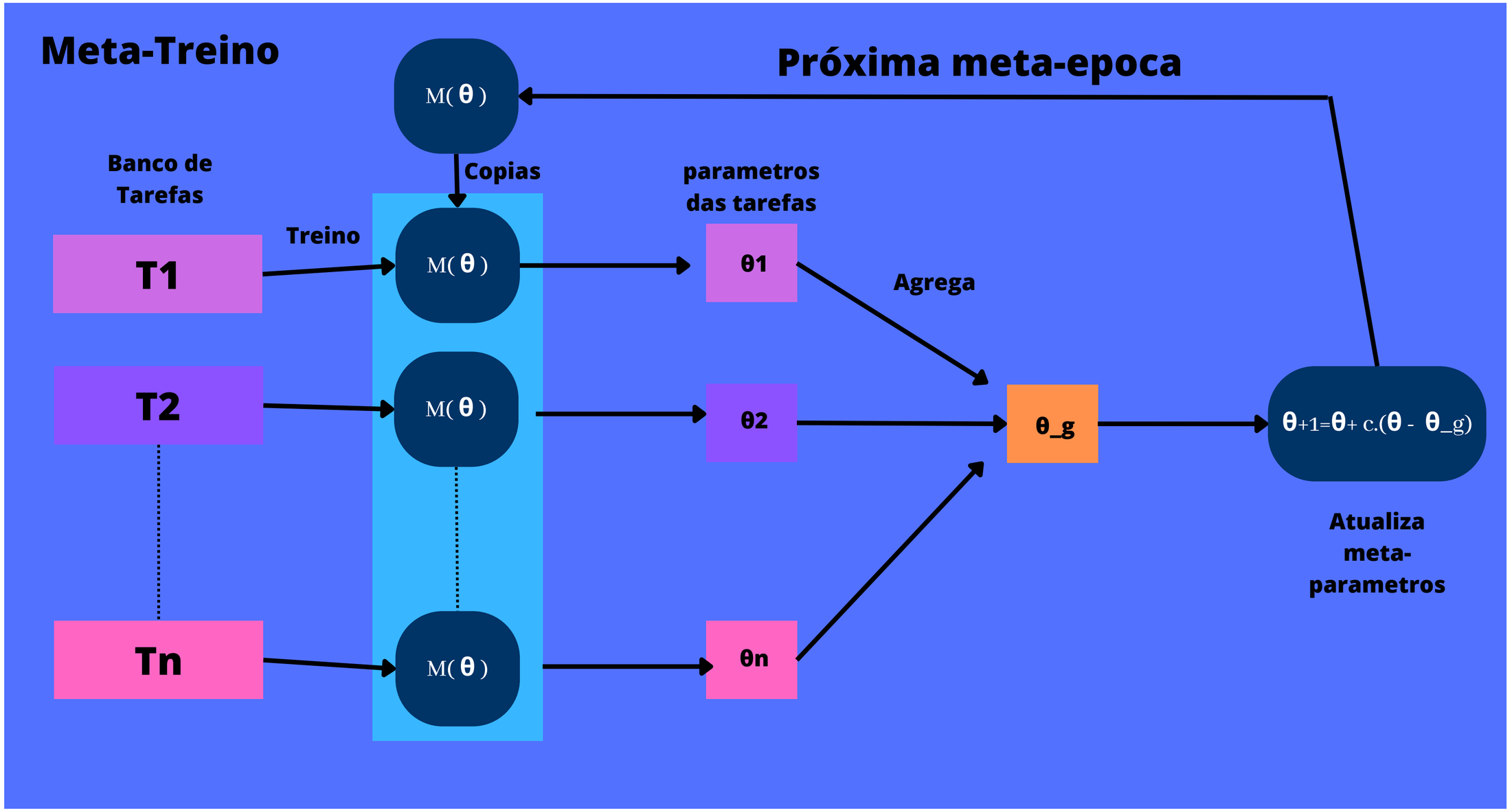


Arara: 1
Beija-Flor: 0

# Few-shot Learning

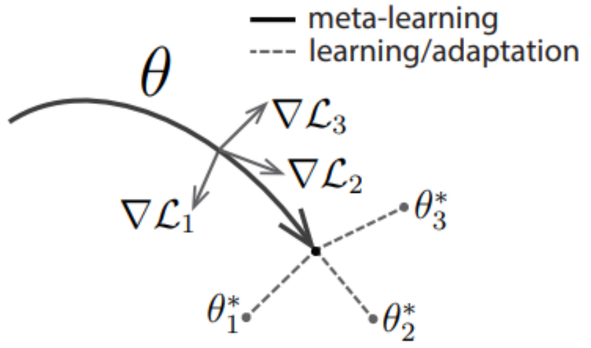

Meta-Learning
Banco de Tarefas
T1
T2
Tn
Meta-Treino
Modelo Pré-Treinado
Inicializar Parametros
Task-Learning
Tn+1
Few-shot Treino
Learner
Classificar
Banco de dados não rotulado

# Meta-Treino


Meta-Treino
Próxima meta-epoca
M( θ )
Banco de Tarefas
Copias
parametros das tarefas
Treino
T1
M( θ )
θ1
Agrega
T2
M( θ )
θ2
θ_g
θ+1=θ+ c.(θ - θ_g)
Atualiza meta-parametros
Tn
M( θ )
θn

# Referências

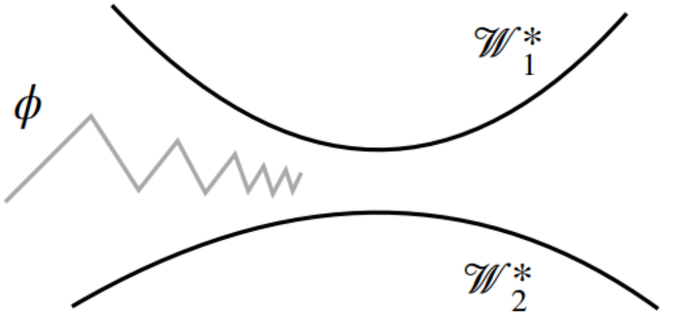
Finn, C., Abbeel, P., and Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. CoRR, abs/1703.03400




$$\min_{\theta} \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta_i'}) = \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_i}(f_\theta)})$$

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_i \sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_i}(f_{\theta_i'})$$

Nichol, A., Achiam, J., and Schulman, J. (2018). On first-order meta-learning algorithms. CoRR, abs/1803.02999




$$\underset{\phi}{\text{minimize}}\, \mathbb{E}_\tau \left[\tfrac{1}{2} D(\phi, \mathcal{W}_\tau^*)^2\right]$$

$$\begin{aligned}\phi &\leftarrow \phi - \epsilon \nabla_\phi \tfrac{1}{2} D(\phi, \mathcal{W}^*)^2 \\ &= \phi - \epsilon(W_\tau^*(\phi) - \phi) \\ &= (1-\epsilon)\phi + \epsilon W_\tau^*(\phi).\end{aligned}$$

$$\begin{aligned}\nabla_\phi \mathbb{E}_\tau \left[\tfrac{1}{2} D(\phi, \mathcal{W}_\tau^*)^2\right] &= \mathbb{E}_\tau \left[\tfrac{1}{2} \nabla_\phi D(\phi, \mathcal{W}_\tau^*)^2\right] \\ &= \mathbb{E}_\tau \left[\phi - W_\tau^*(\phi)\right], \text{where } W_\tau^*(\phi) = \arg\min_{W \in \mathcal{W}_\tau^*} D(W, \phi)\end{aligned}$$