Preliminary Work

MAML

MAML

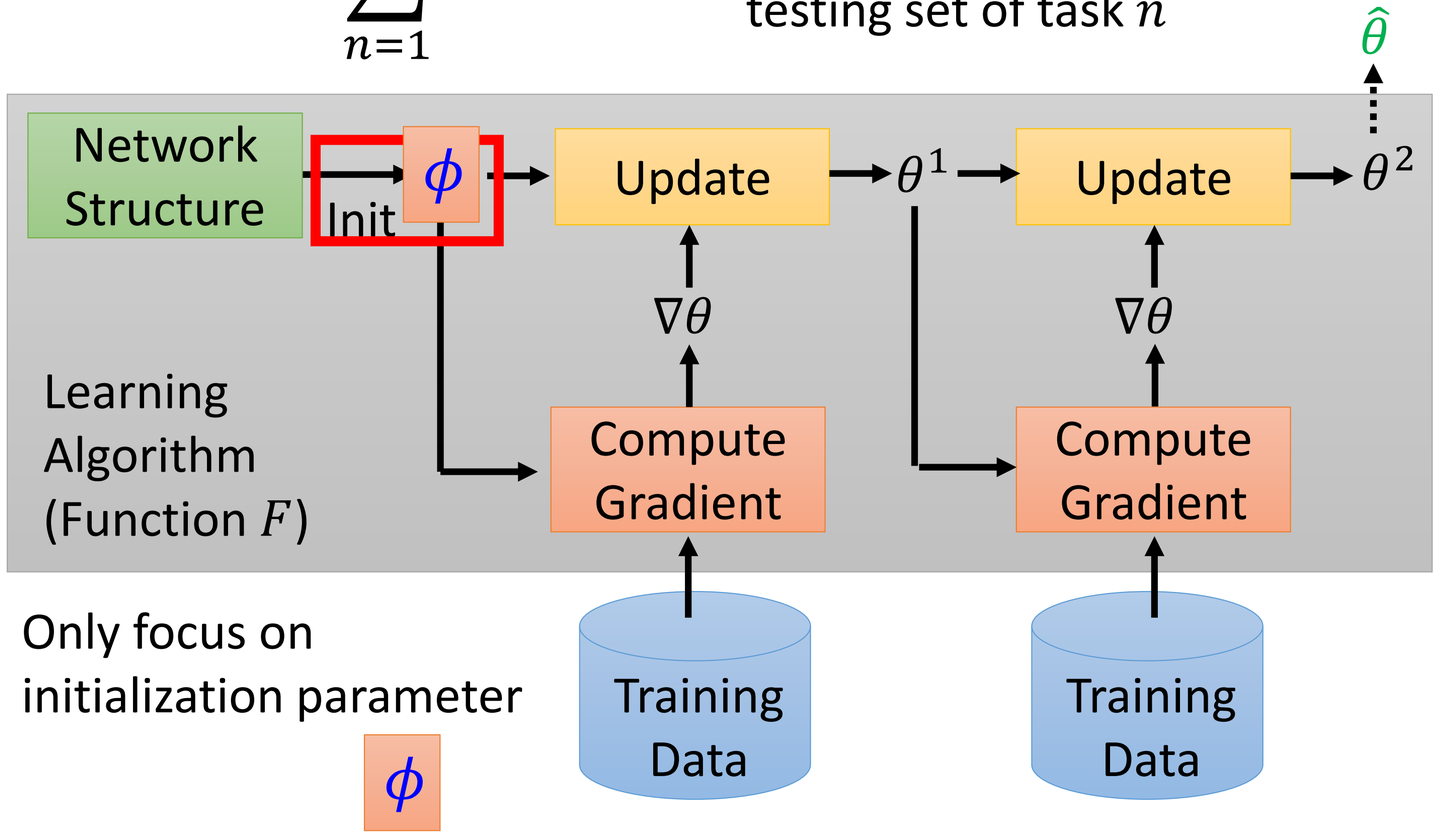
$\hat{\theta}^n$: model learned from task n

$\hat{\theta}^n$ depends on ϕ

Loss Function:

$$L(\phi) = \sum_{n=1}^{N} l^n(\hat{\theta}^n)$$

$l^n(\hat{\theta}^n)$: loss of task n on the testing set of task n



https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Meta1%20(v6).pdf

Preliminary Work

MAML

MAML

Loss Function:

$$L(\phi) = \sum_{n=1}^{N} l^n\left(\hat{\theta}^n\right)$$

$\hat{\theta}^n$: model learned from task n

$\hat{\theta}^n$ depends on ϕ

$l^n\left(\hat{\theta}^n\right)$: loss of task n on the testing set of task n

How to minimize $L(\phi)$? Gradient Descent

$$\phi \leftarrow \phi - \eta \nabla_{\phi} L(\phi)$$

Model Pre-training

Widely used in transfer learning

Loss Function:

$$L(\phi) = \sum_{n=1}^{N} l^n(\phi)$$

https://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Meta1%20(v6).pdf