

1

1) Task  $(D_{\text{train}}, D_{\text{test}})$   
 $\{x_{\text{train}}, y_{\text{train}}\}$   $\{x_{\text{test}}, y_{\text{test}}\}$

2) Goal is to learn how to do well on Test  
Meta-learning directly minimize  $\text{loss}(f, \underline{D_{\text{test}}} | D_{\text{train}})$   
across many tasks  
so for one task  $f(x | D_{\text{train}})$   
 $\rightarrow$  does well

3) MAML (optimization based meta-L) <sup>on  $D_{\text{test}}$</sup>

$f \rightarrow$  taking one (or more) gradient steps  
on  $\mathcal{L}(D_{\text{train}})$  for some nn  $\hat{f}$

$$\nabla_w \mathcal{L} \left( \underbrace{g_w \rightarrow \mathcal{L}_{\text{loss}}(g_w(D_{\text{train}}))}_{\phi} (D_{\text{test}}) \right)$$



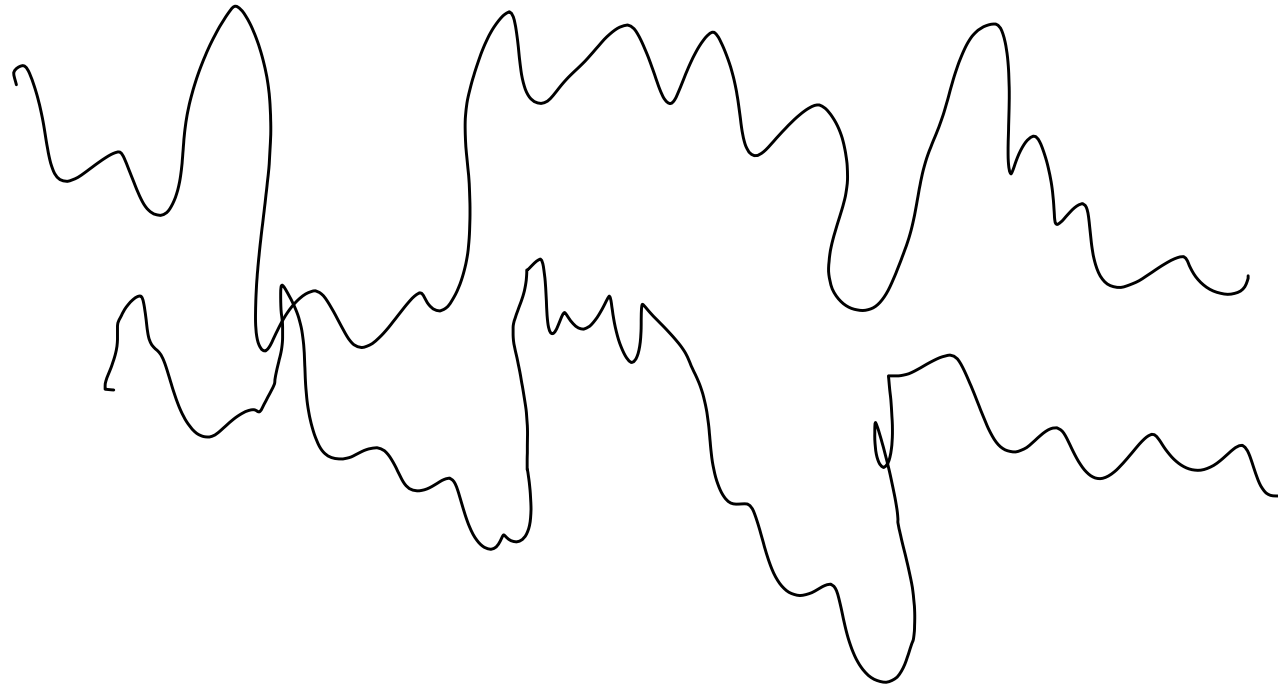
- K Y 000 2 0

+ + + 2  
K-1 + + }  
\_\_\_\_\_ 0

HW

- Run MAML
- ① different fcs
  - ② different shots / ways
  - ③ different degrees of dan-overlap
  - ④ rate of convergence  
initia train is much less  
accurate

Exi



A

MAMU .

→ 2nd order diff

→ Task - far shot  
Class over

→ far



Thing

Met

Shots, ways often                      Train.

HW

for  
class overfit

training regu  
spat.

Model > Answer  
Ts. =

And

pretraining is MAKL

only on test data

no adaptation

have reliable  
=