7.6

Leveraging Pre-trained Models with Transfer Learning

Part 2: Different Transfer Learning Variants

Sebastian Raschka and the Lightning AI Team

Transfer learning variants

Dataset perspective

1

Step 1: train on ...

Large general dataset

Step 2: fine-tune on ...

Smaller target dataset

1

Step 1: train on ...

Large general dataset

Step 2: fine-tune on ...

Smaller target dataset

2

Step 1: train on ...

Large general dataset + Smaller target dataset

1

Step 1: train on ...

Large general dataset

Step 2: fine-tune on ...

Smaller target dataset

2

Step 1: train on ...

Large general dataset + Smaller target dataset

3

Step 1: train on ...

Large general dataset

Step 2: train on ...

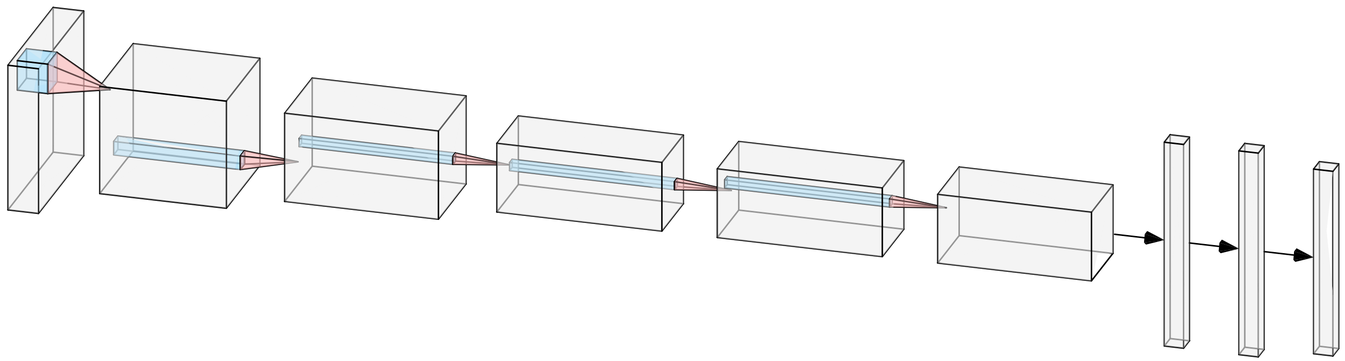
Large general dataset + Smaller target dataset

Transfer learning variants

Model perspective

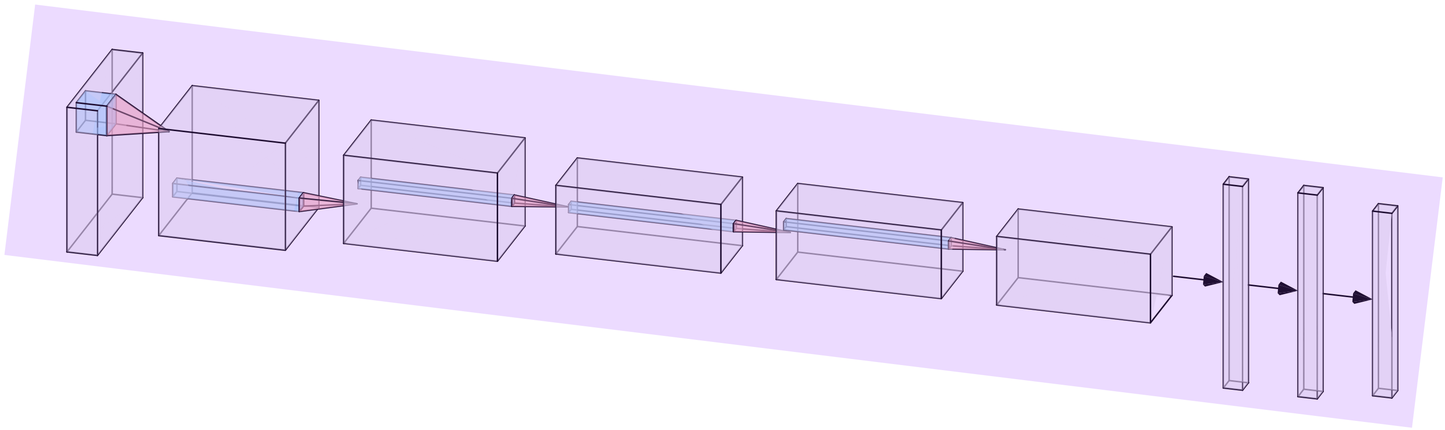
❶ Fine-tune last layer

Step 1: train whole model on large dataset

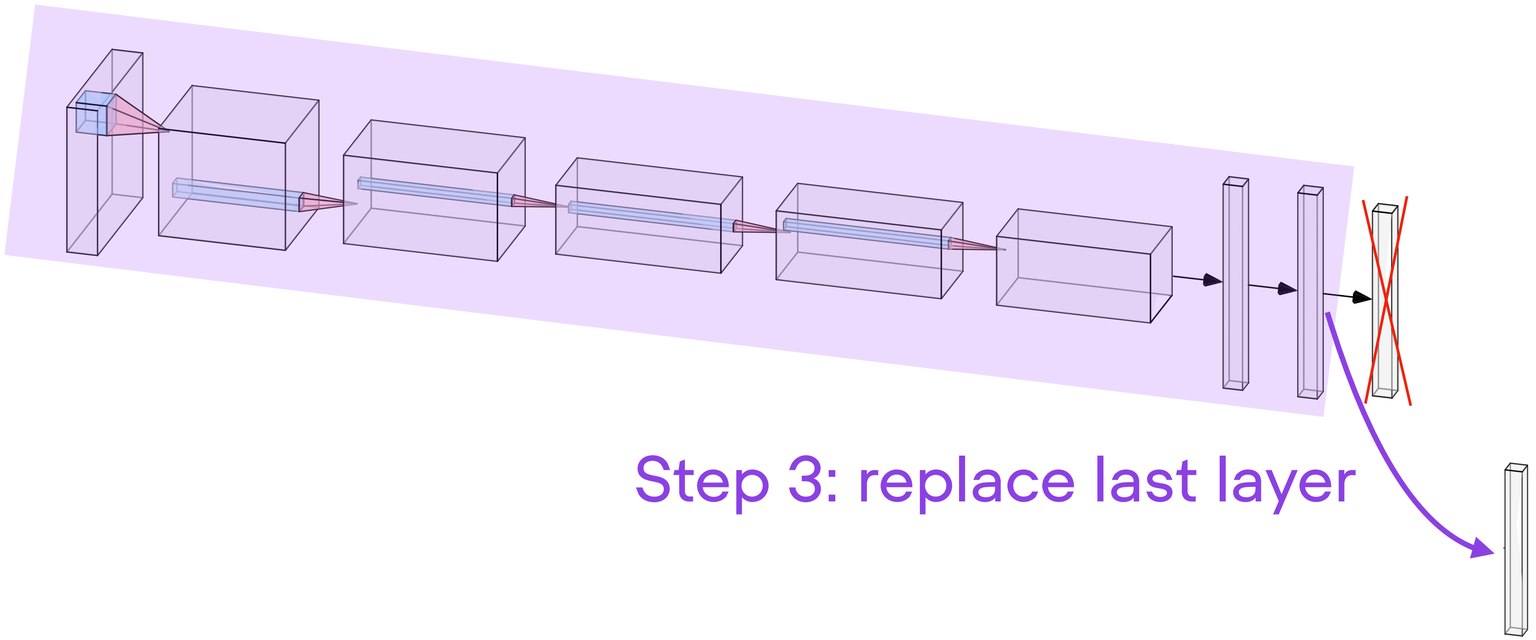


1 Fine-tune last layer

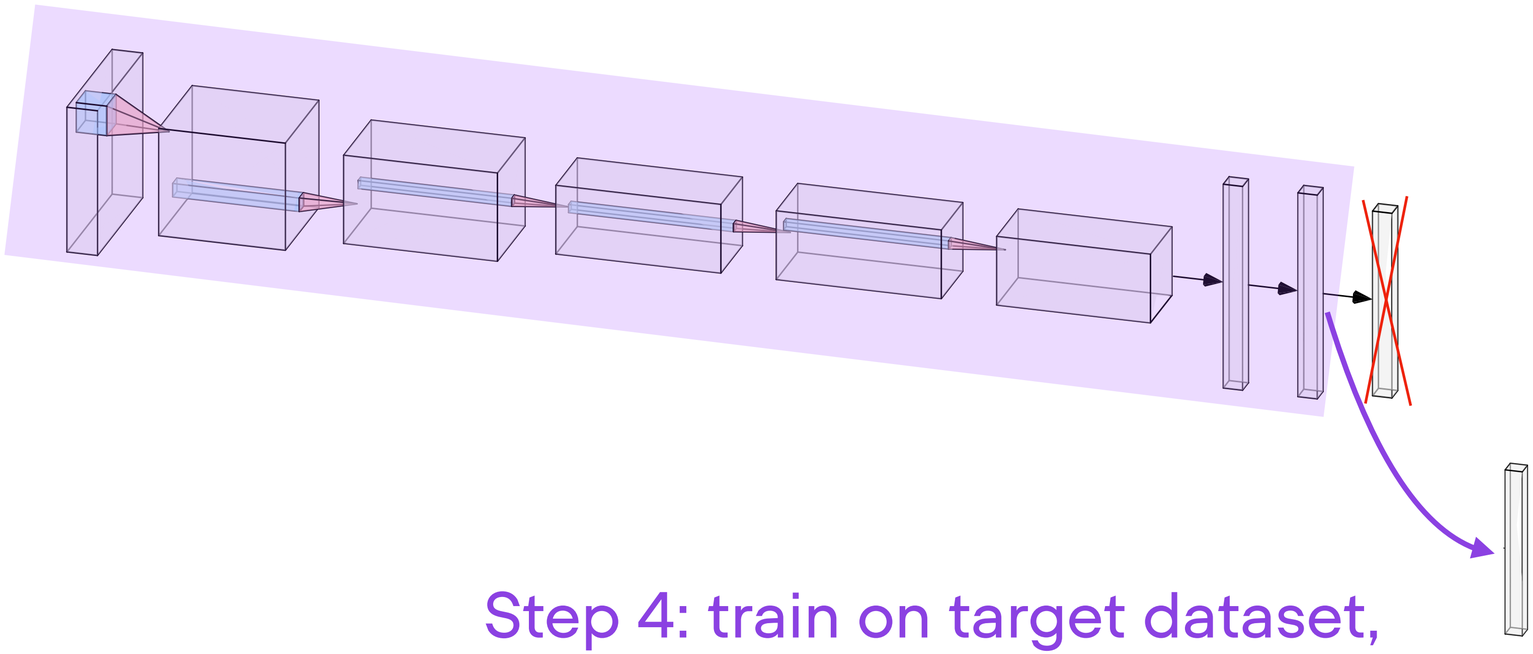
Step 2: freeze weights



❶ Fine-tune last layer



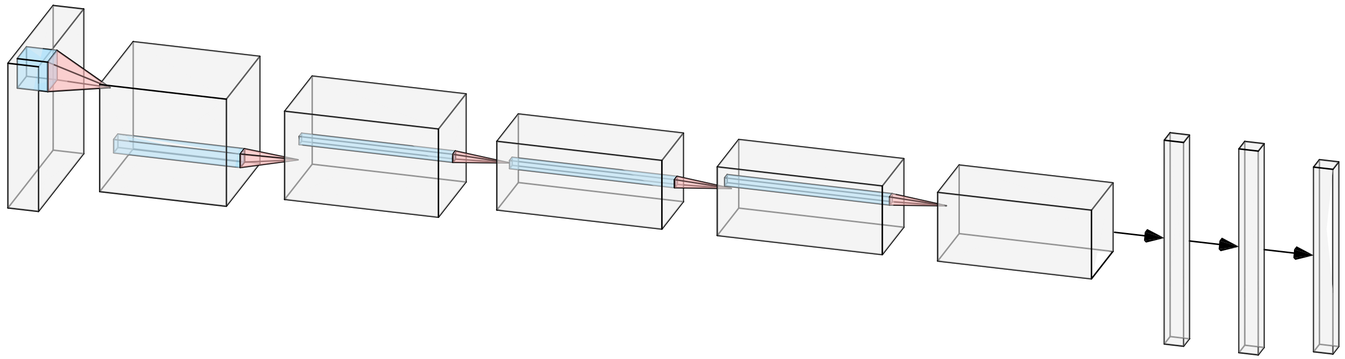
❶ Fine-tune last layer



Step 4: train on target dataset,
but only update new output layer

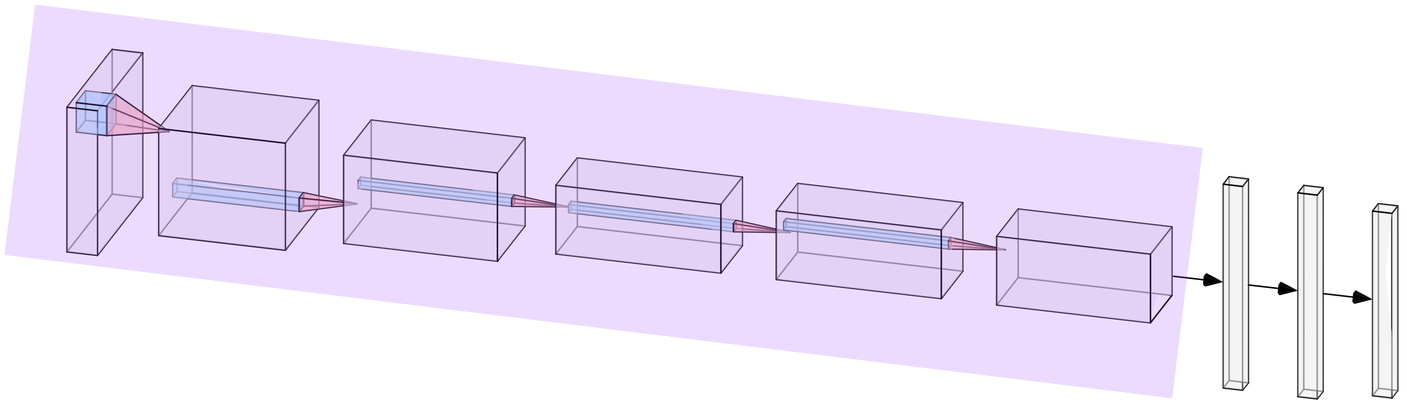
❷ Fine-tune last layers

Step 1: train whole model on large dataset

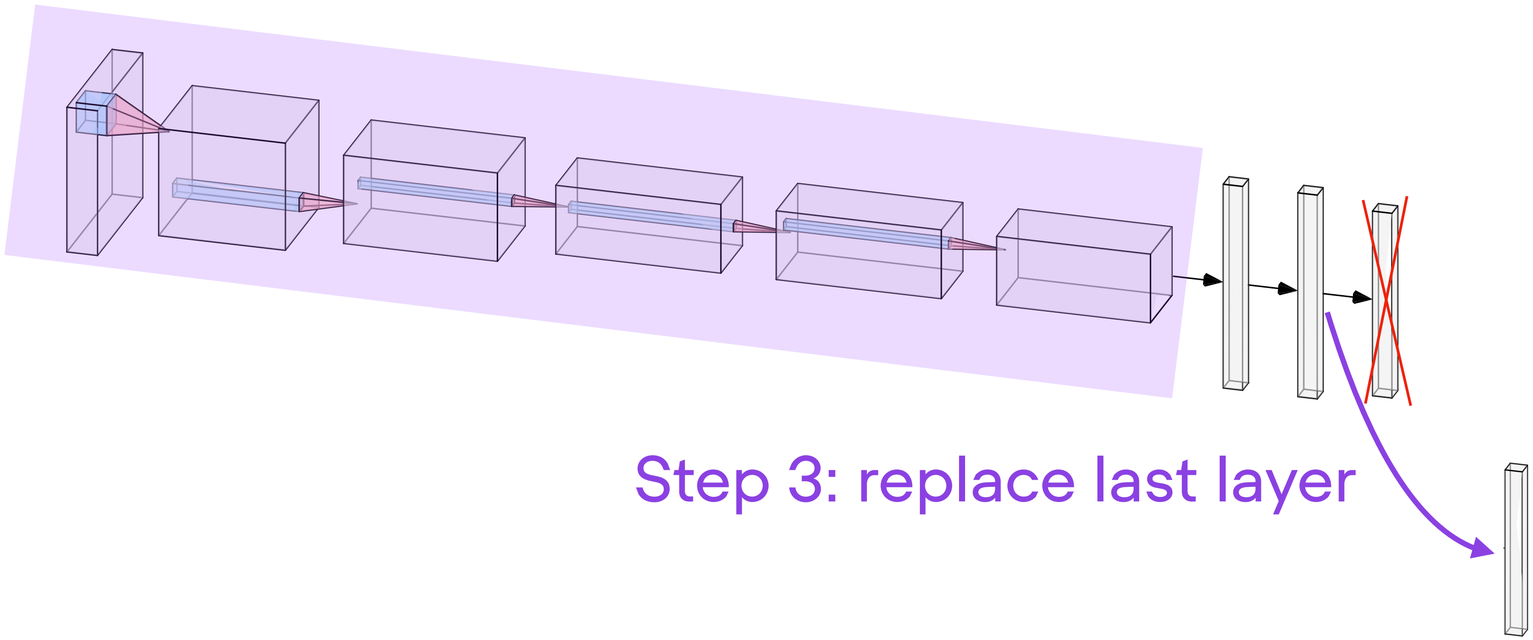


❷ Fine-tune last layers

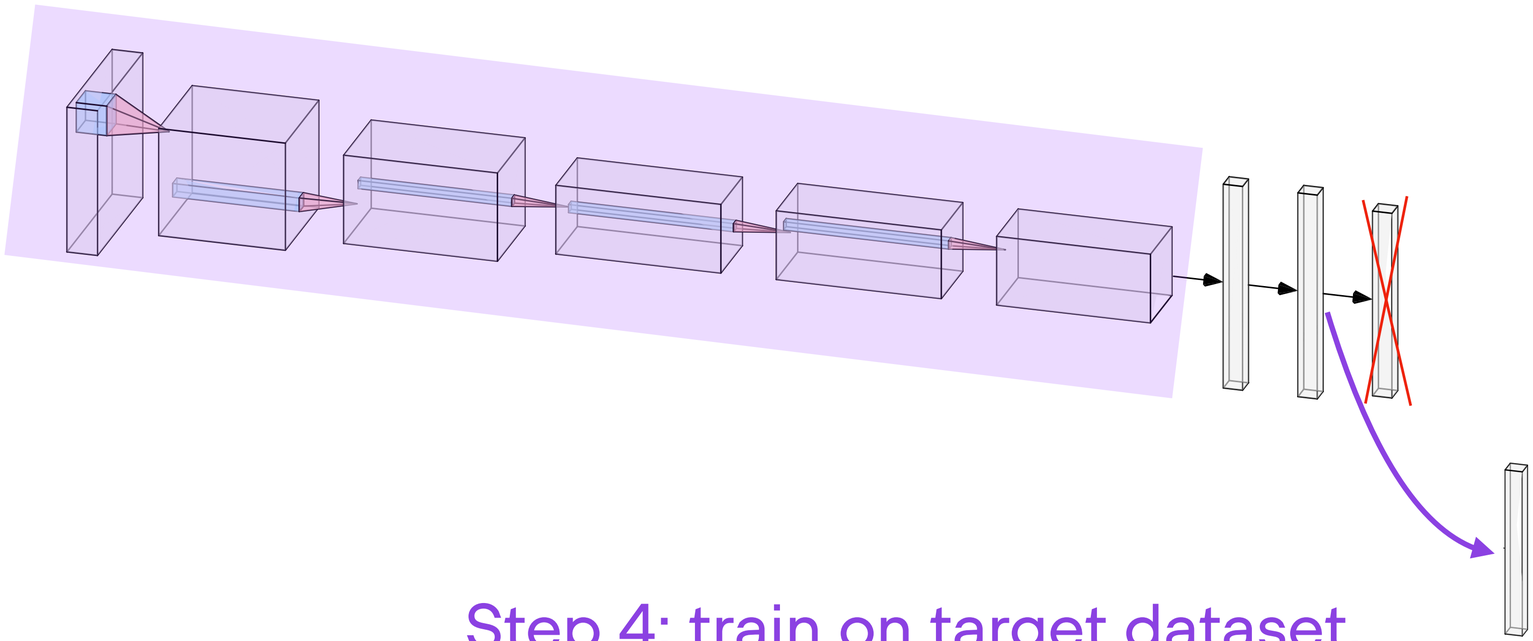
Step 2: freeze weights



❷ Fine-tune last layers



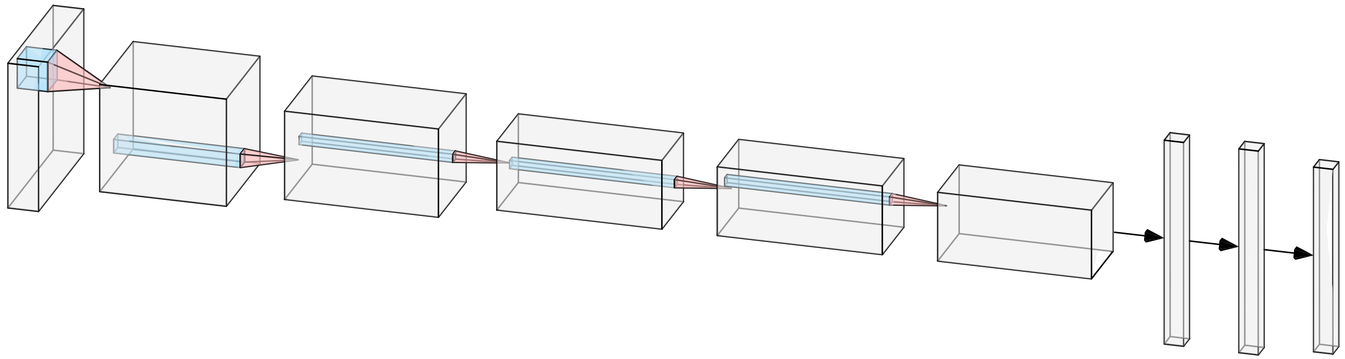
❷ Fine-tune last layers



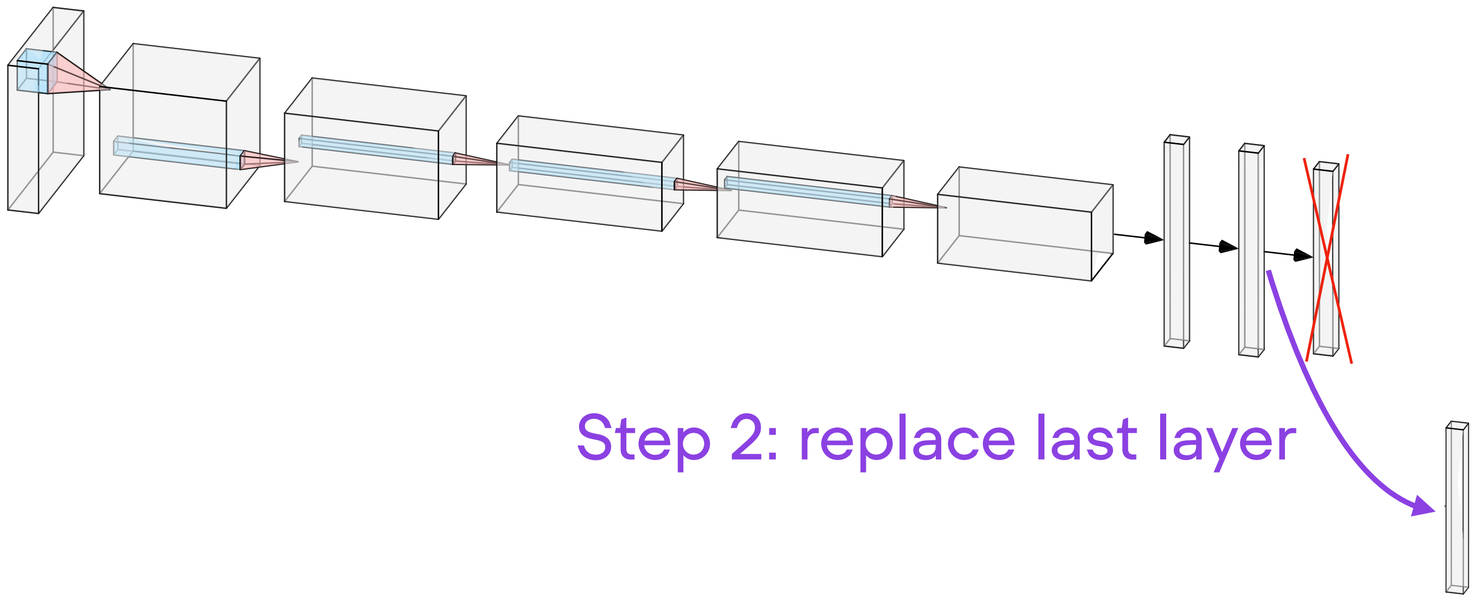
Step 4: train on target dataset,
but only update the last output layers

3 Fine-tune the whole model

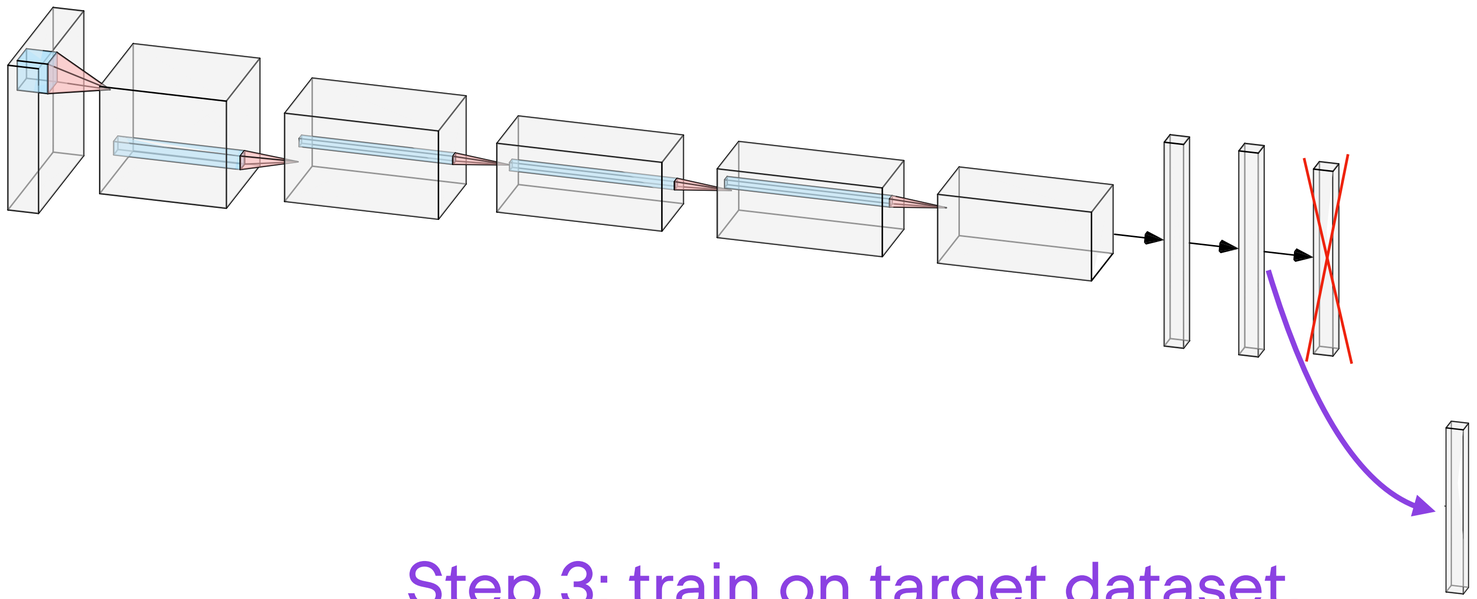
Step 1: train whole model on large dataset



3 Fine-tune the whole model



3 Fine-tune the whole model



Step 3: train on target dataset,
but only update the last output layers

Which paradigm to use is situation-dependent
(~try it like a hyperparameter)

Next: Let's try it out in practice