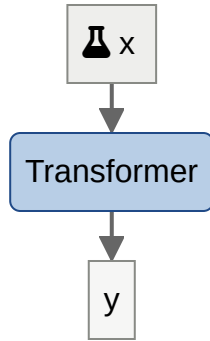
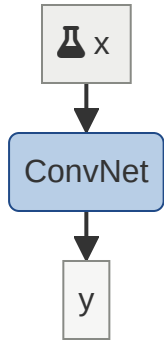


Making It Work

Recap

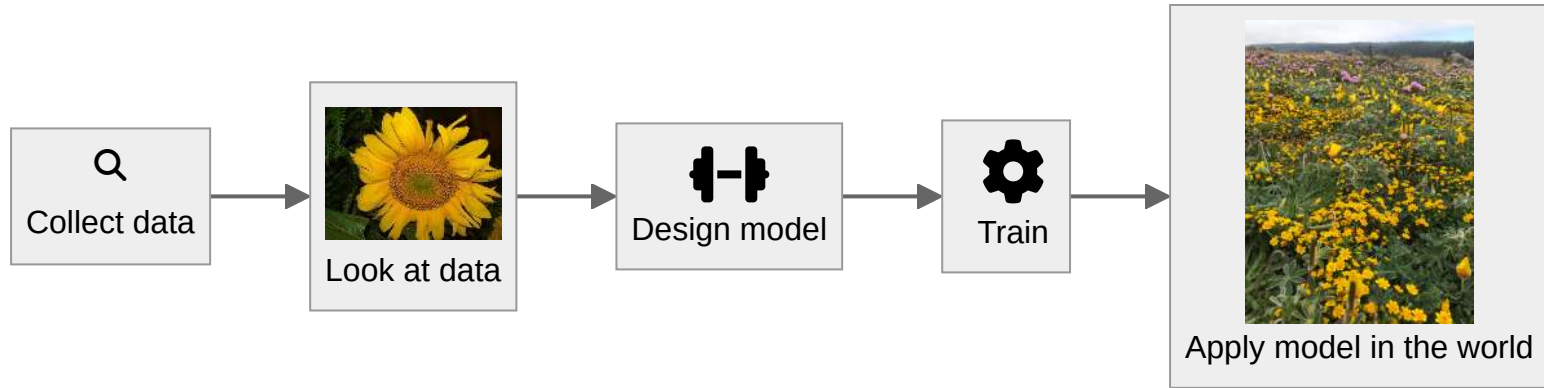
Models



Optimization

```
m = 0
for epoch in range(n):
    for (x, y) in dataset:
        J = ∇l(θ|x,y)
        m = J + momentum * m
        θ = θ - ε * m.mT
```

How to Tackle a New Problem

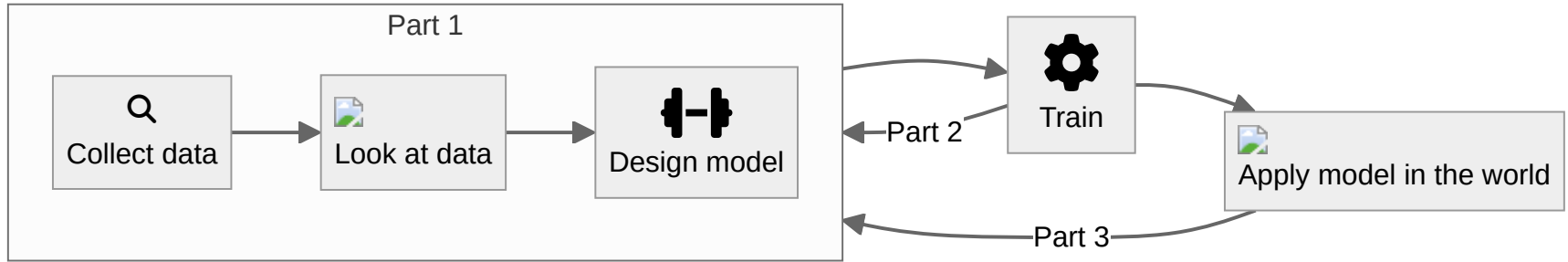


When Will This Work?

- Never! 😞

Making It Work

In practice, this is an iterative process



Making It Work

Part 1: Data and Advanced Network Design

Part 2: Advanced Training

Part 3: Overfitting

