⟨ Bias / Variance and neural network ⟩

- Large neural networks are low bias machines

```
┌─────────────────┐              ┌──────────────────────┐
│ Does it do well │    Yes       │ Does it do well      │   Yes
│ on the train set?│ ──────────→ │ on the validation set?│ ────────→  Done !
│        Jtrain    │              │           Jcv         │
└─────────────────┘              └──────────────────────┘
         │                                    │
         │ No (high bias)                     │ No (high variance)
         ↓                                    ↓
    Bigger network                       Get more data
```

until it does
well on train set
(archeives level   (more layers
of error on   more units)
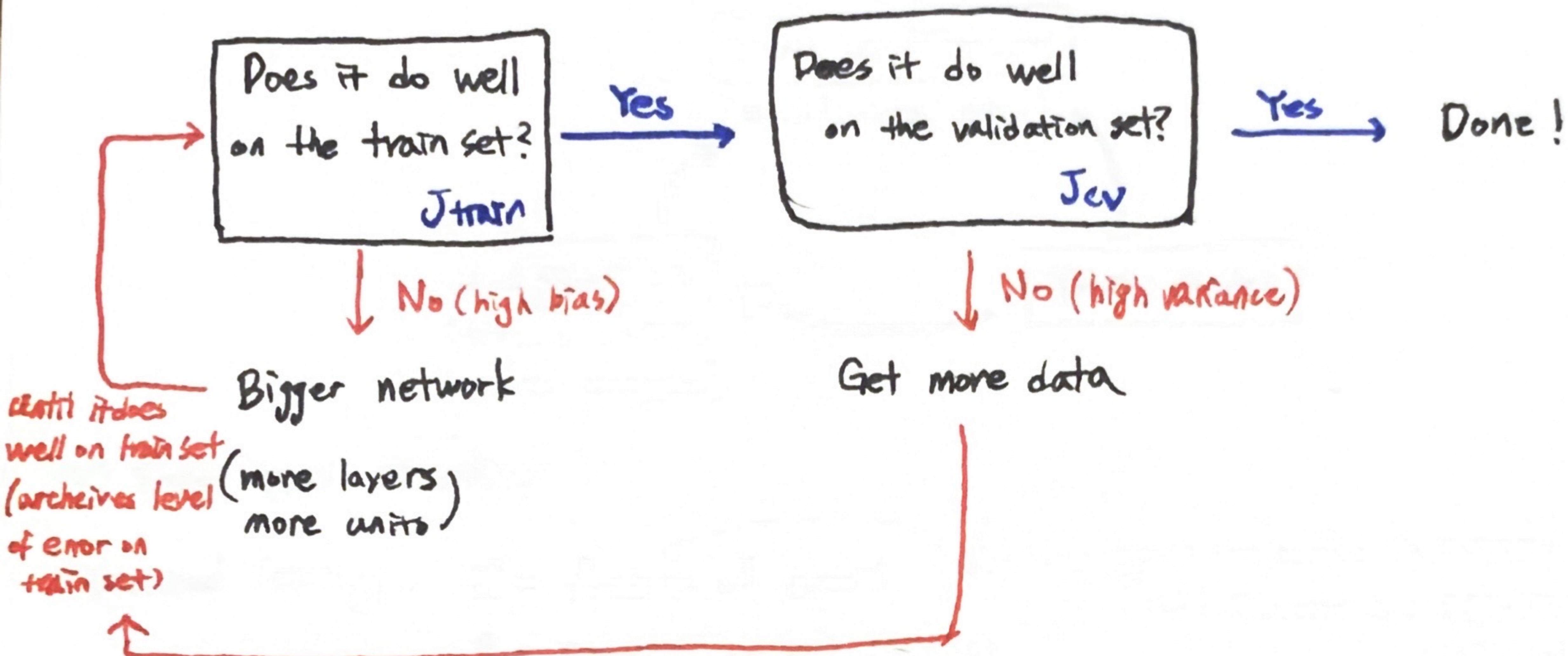train set)

- Large neural network will usually do as well or better than a smaller one
  trade off → so long as regularization is chosen on the larger neural network appropriately
    - Computationally expensive
    - slow training time