Machine Learning (ML)

- Supervised (regression, classification)
  - Labeled Data
- Unsupervised (clustering, PCA)
- Reinforcement
  - Delayed Labels
Normal Program

input → Program → results

Parametrized Program

inputs → Program → results

Learned Program

inputs → Architecture → predictions
parameters

Training Process

labels → Evaluation → loss

How good/bad are the predictions?
Inference Process

inputs (unseen) $\rightarrow$ Architecture $\rightarrow$ results

parameters (good values)

---

size

zip

age

# beds

# baths

pool?

Walkability

weather

family size

Feature Engineers (not for NN)

price
2,100 sq. ft $\rightarrow 0.3$

$\text{size}$

$91711$ $\rightarrow$ $-1$ to $1$

$\text{zip}$

$98$ $\rightarrow$ $-1$ to $1$

$\text{age}$

$\# \text{ beds}$

$\# \text{ baths}$

pool?

$2013$

NN

$\text{price}$

1,000,100

image

NN

0.0/3
1.0/3
2.0/3

left
straight
right
Limitations

- NN are often called "Black Box" models
- Can only train if we have labeled data
- They only do what you tell them to
- Just make predictions
- Easy to crack a bad model (biased)