Module 1.0 - Mini-ML

Model

- Models: parameterized functions.
    - $m(x; \theta)$
    - x - input
    - m - model
- Initial Focus:
    - θ - parameters

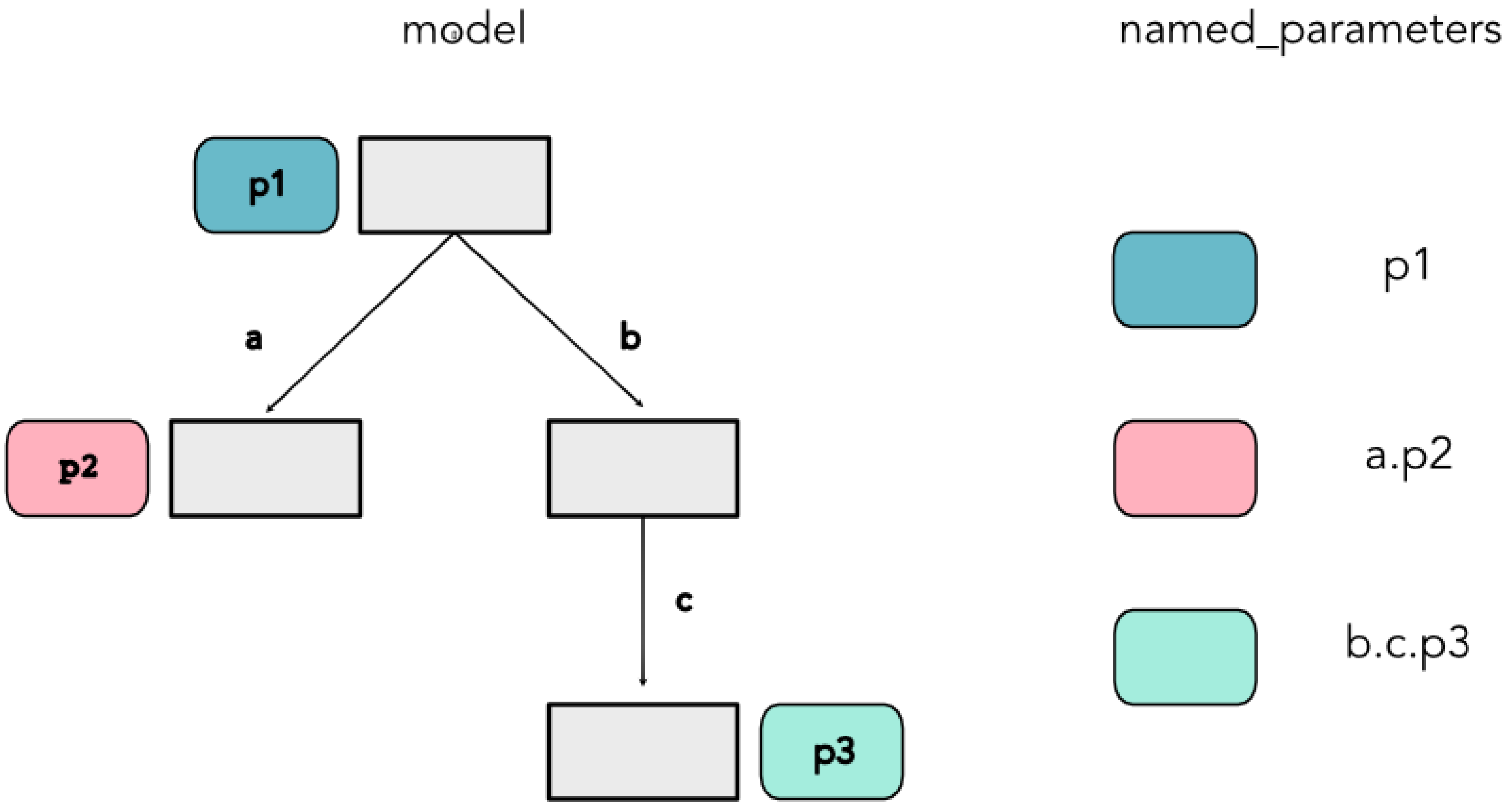
Specifying Parameters

- Datastructures to specify parameters
- Requirements
  - Independent of implementation
  - Compositional

Module Example

```python
from minitorch import Module, Parameter


class OtherModule(Module):
    pass


class MyModule(Module):
    def __init__(self):
        # Must initialize the super class!
        super().__init__()

        # Type 1, a parameter.
        self.parameter1 = Parameter(15)

        # Type 2, user data
        self.data = 25

        # Type 3. another Module
        self.sub_module = OtherModule()
```

Module Naming



Lecture Quiz

Quiz

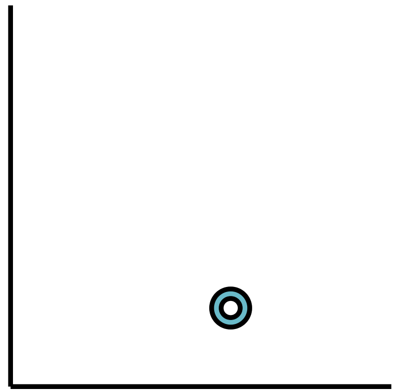
Outline

- Model
- Parameters
- Loss

Datasets

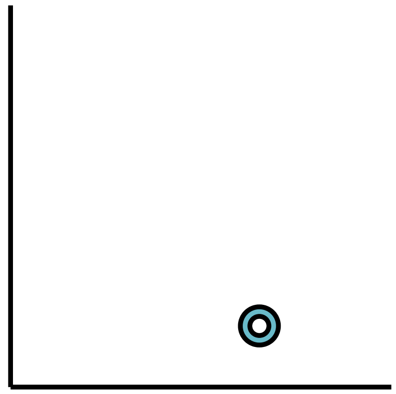
Data Points

- Convention x



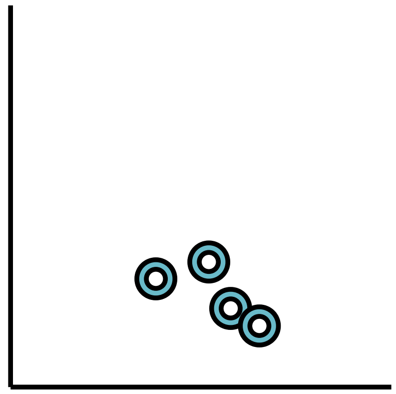
Data Points

- Convention x



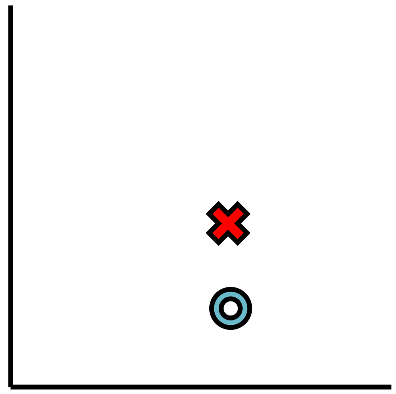
Data Points

- Convention x



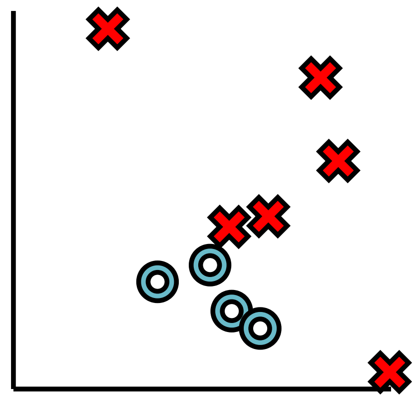
Data Labels

- Convention y



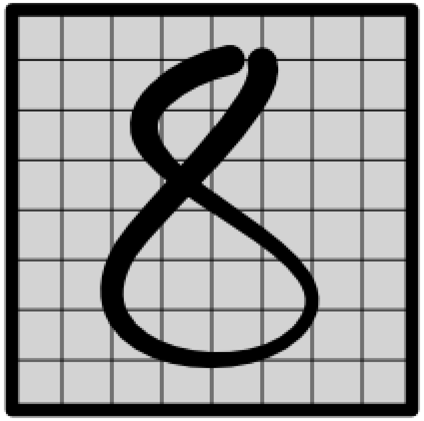
Training Data

- Set of datapoints, each (x, y)



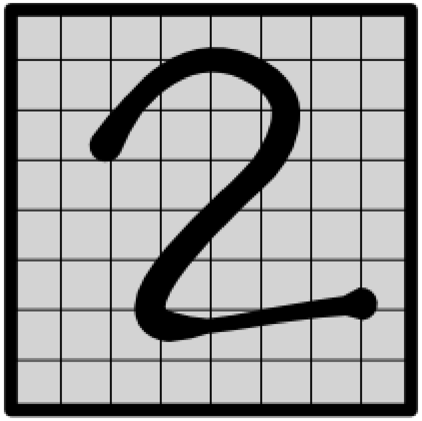
Data Points

- Convention x

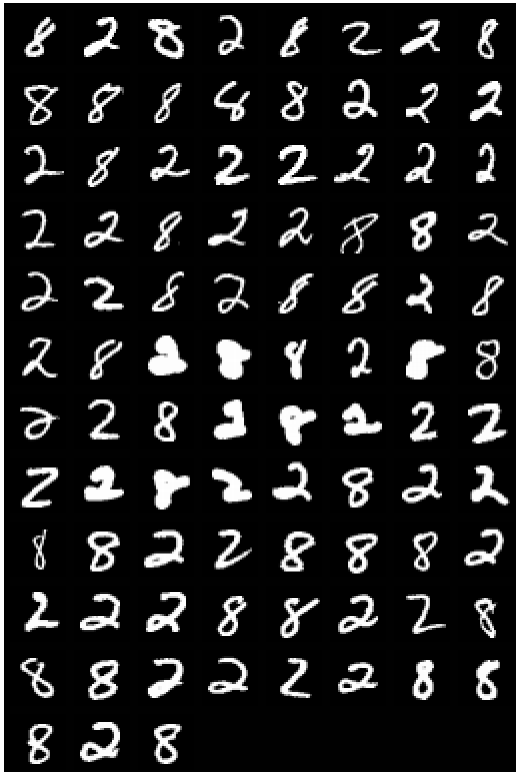


Data Points

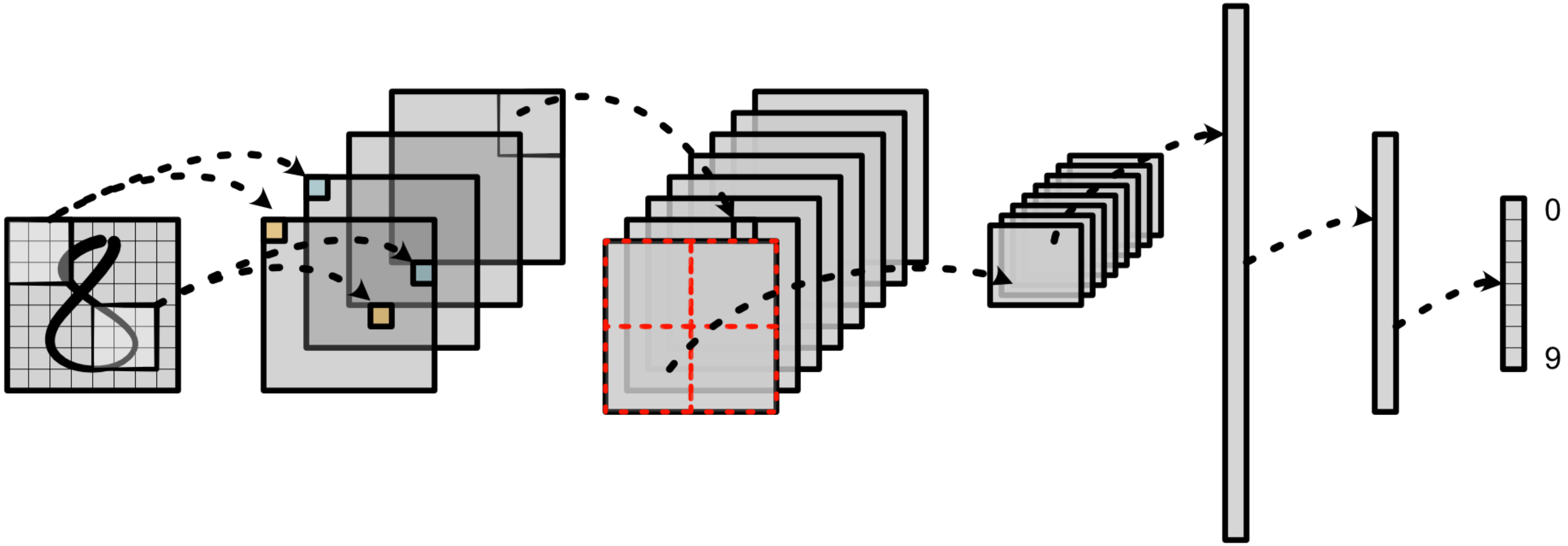
- Convention x



Data Set

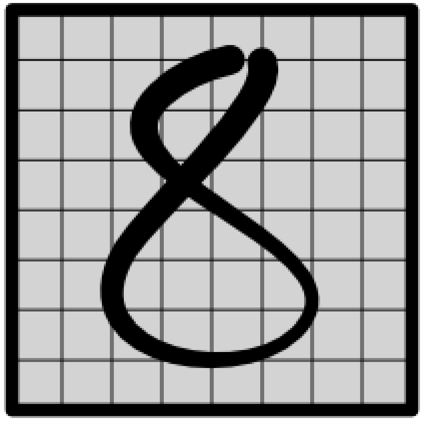


Network



Data Labels

- Convention y



Model

Models

- Functions from data points to labels
- Functions $m(x; \theta)$
- Any function is okay (e.g. Modules)

Example Model

- Example of a simple model

  x = (0.5, 0.2)

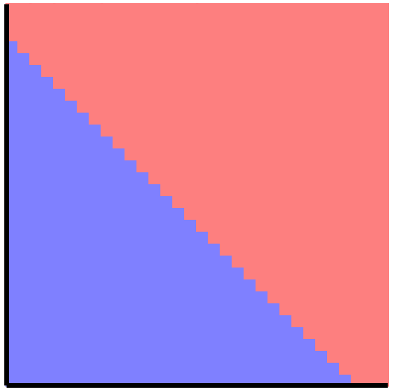
```python
class Model:
    def forward(self, x):
        return 0 if x[0] < 0.5 else 1
```

Model 1

- Linear Model

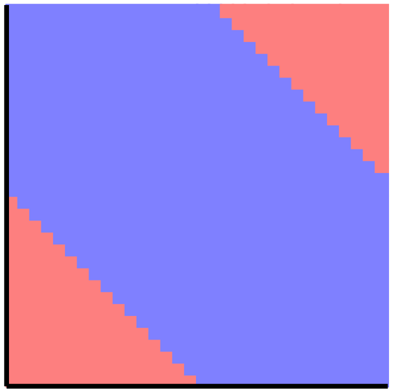
```
@dataclass
class Linear:
    # Parameters
    w1: float
    w2: float
    b: float

    def forward(self, x1: float, x2: float) -> float:
        return self.w1 * x1 + self.w2 * x2 + self.b
```

Model 1



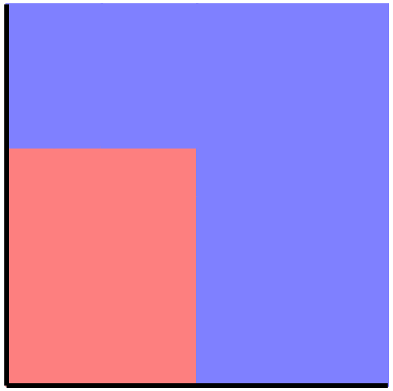
Model 2

```
@dataclass
class Split:
    # Submodules
    m1: Linear
    m2: Linear

    def forward(self, x1, x2):
        return self.m1.forward(x1, x2) * self.m2.forward(x1, x2)
```



Model 3

```
class Part:
    def forward(self, x1, x2):
        return 1 if (0.0 <= x1 < 0.5 and 0.0 <= x2 < 0.6) else 0
```



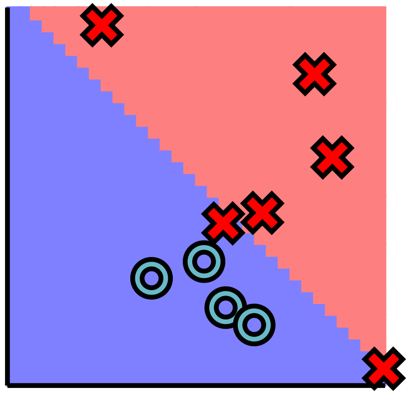
Parameters

Parameters

- Knobs that control the model
- Any information that controls the model shape

Parameters

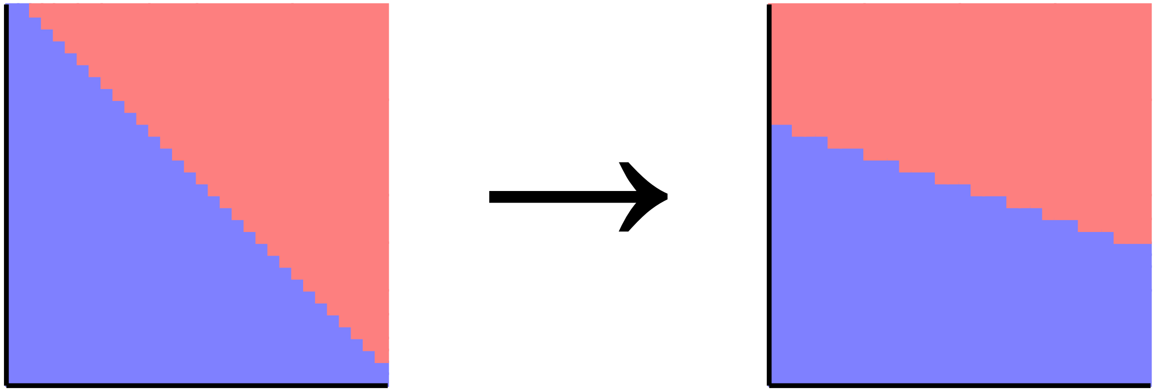
- Change θ



Linear Parameters

a. rotating the linear separator

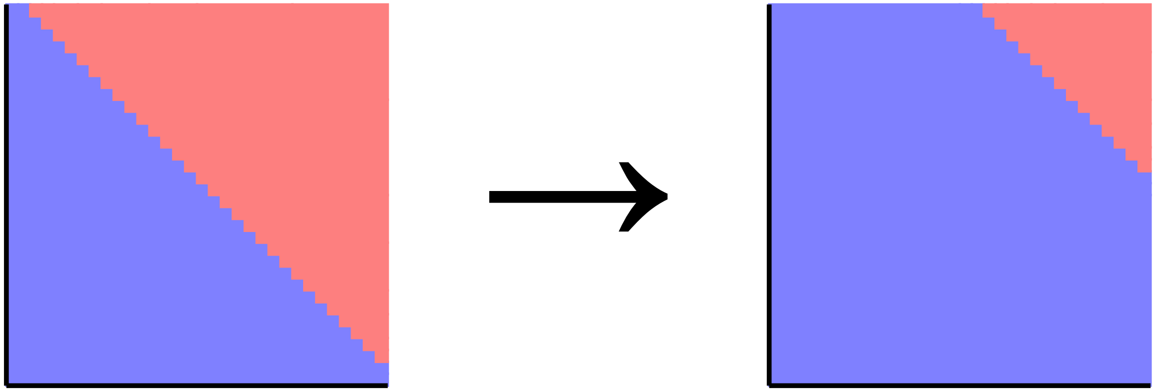
```
model1 = Linear(1, 1, -1.0)
model2 = Linear(0.5, 1.5, -1.0)
```



Linear Parameters

b. changing the separator cutoff

```
model1 = Linear(1, 1, -1.0)
model2 = Linear(1, 1, -1.5)
```



Math

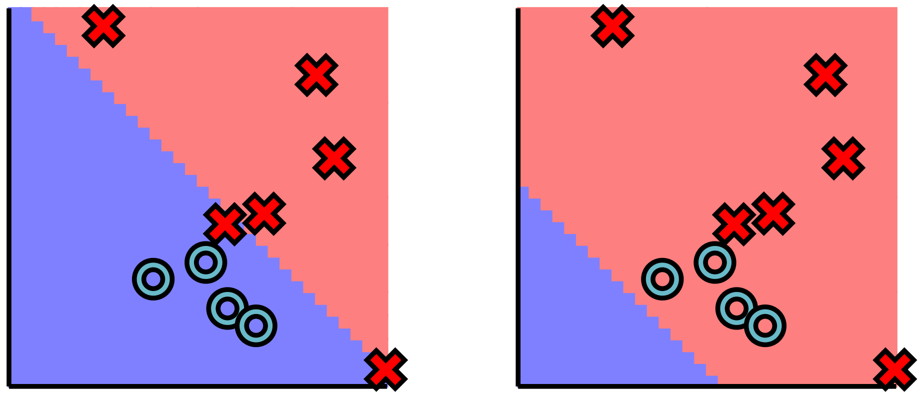
- Linear Model

$$m(x; w, b) = x_1 \times w_1 + x_2 \times w_2 + b$$

```
def forward(self, x1: float, x2: float) -> float:
    return self.w1 * x1 + self.w2 * x2 + self.b
```

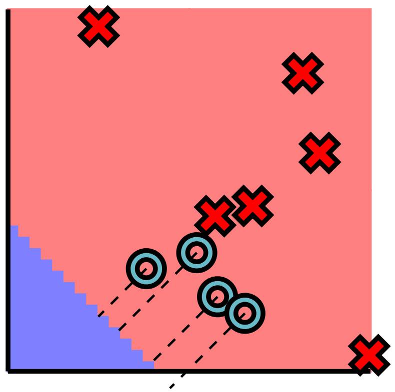
Loss

What is a good model?

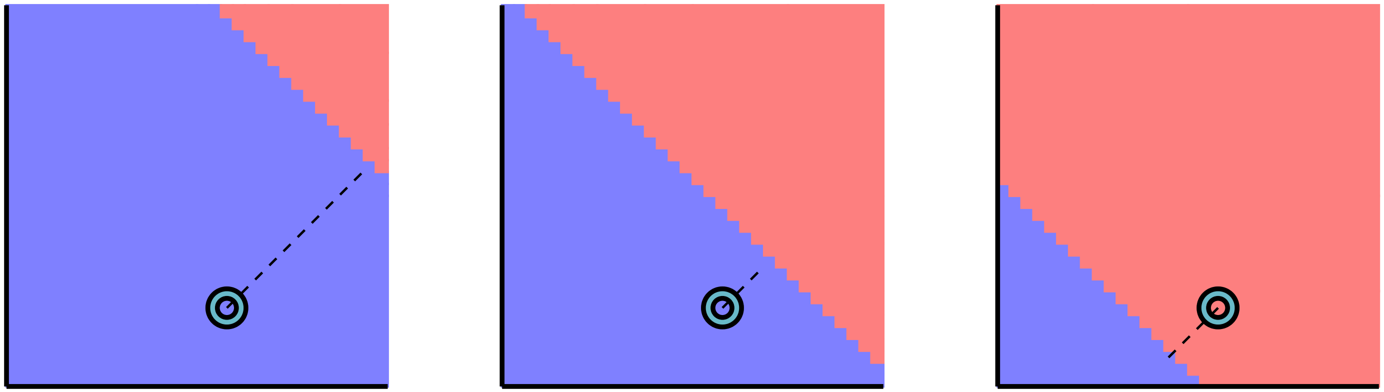


Distance

- $|m(x)|$ correct or incorrect



Points

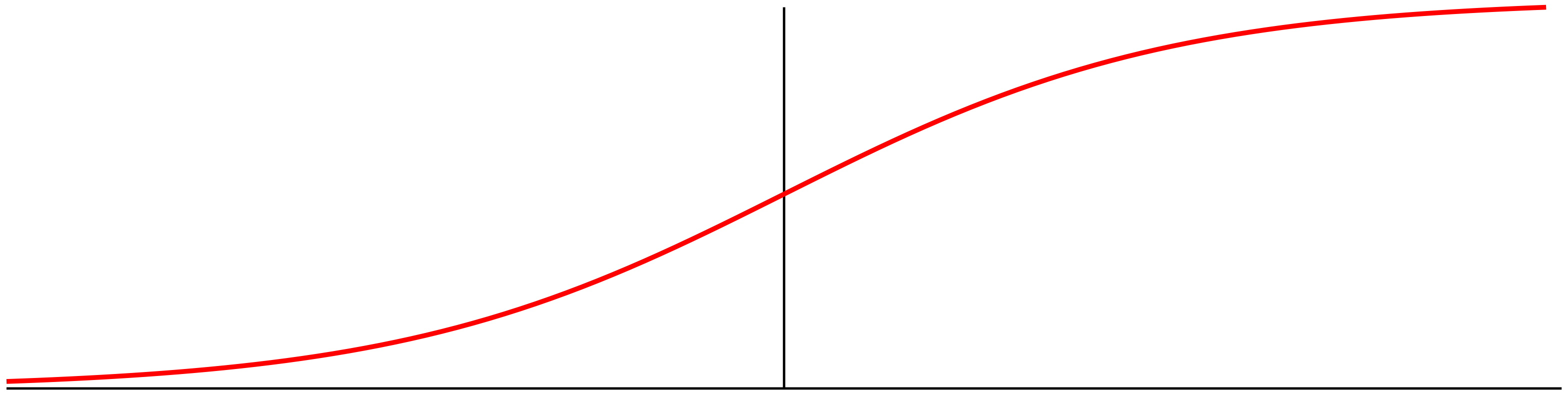


Loss

- Loss weights our incorrect points
- Uses distance from boundary

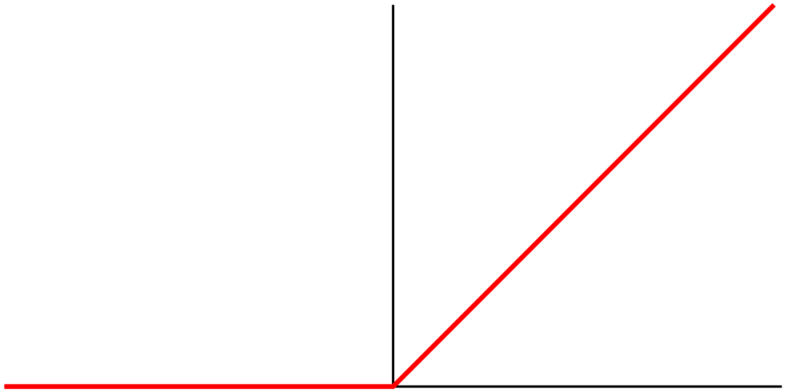
  $L(w_1, w_2, b)$ is loss, function of parameters.

Sigmoid Function

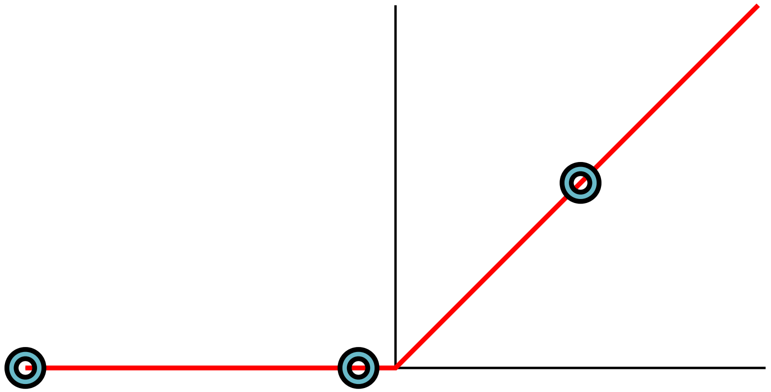


ReLU loss

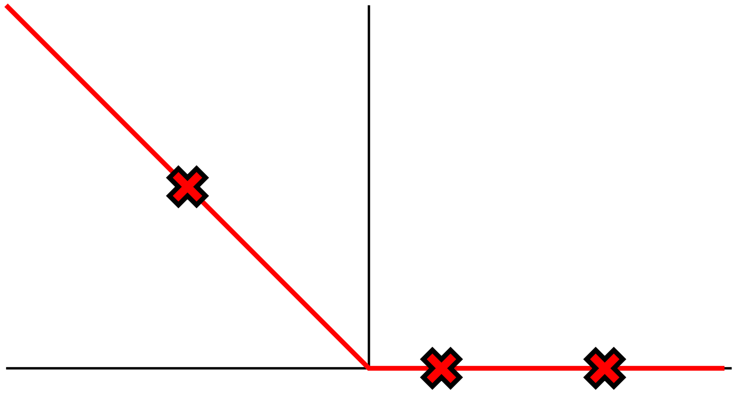
```
def point_loss(m_x):
    return minitorch.operators.relu(m_x)
```



Loss of points



Loss of points



Full Loss

```python
def full_loss(m):
    l = 0
    for x, y in zip(s.X, s.y):
        l += point_loss(-y * m.forward(*x))
    return -l
```

Playground

Playground

Q&A

```python
import torch
```

```python
x = torch.tensor([0.0], requires_grad=True)
torch.relu(x).backward()
x.grad
```

```
tensor([0.])
```