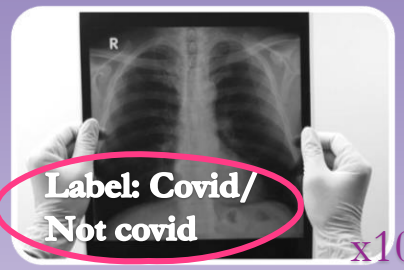
MLDL-I
Machine Learning and Deep Learning - I
Lecture 2

Supervised Learning: An Overview

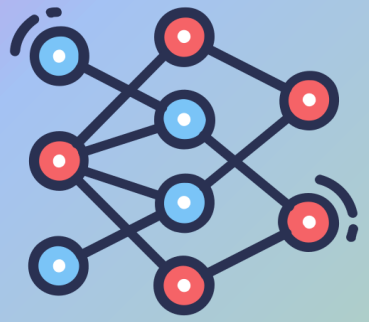
Data



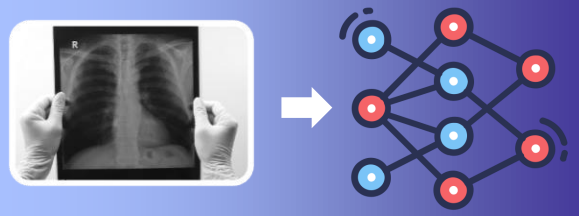
x10000

Rooms	Bathroom	BuildingArea	Price
2	1.0	NaN	1480000.0
2	1.0	79.0	1035000.0
3	2.0	150.0	1465000.0
3	2.0	NaN	850000.0

Training



Inference



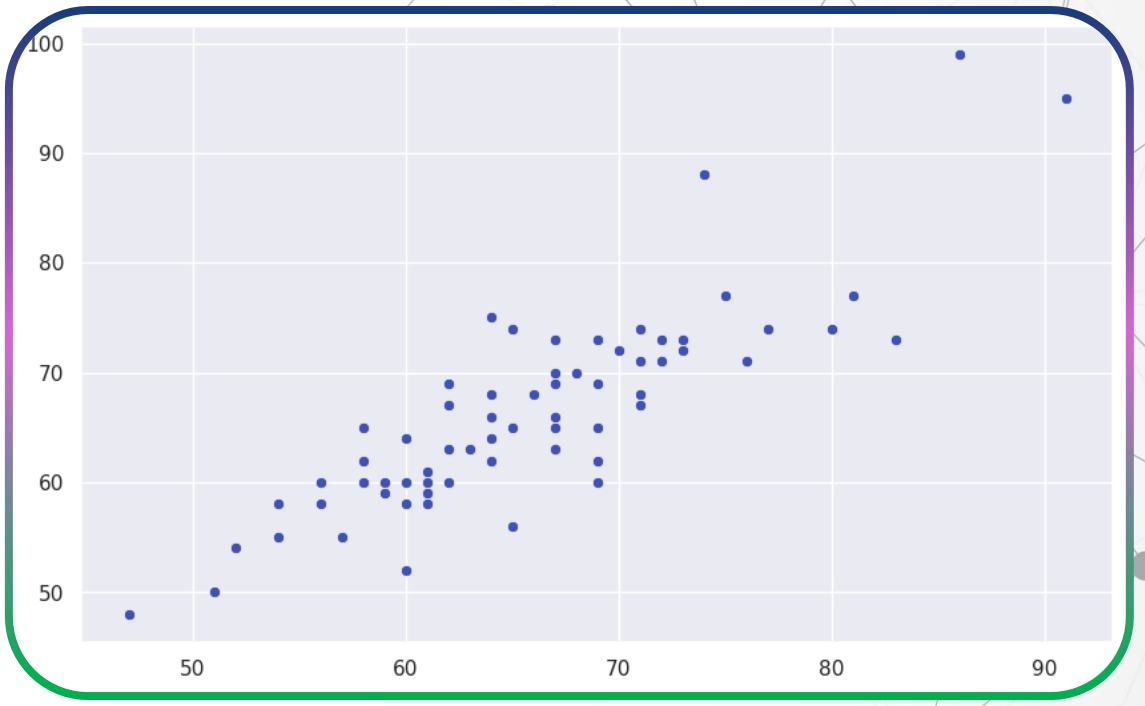
Covid	Not Covid
6% Probability	94% Probability

Rooms	Bathroom	BuildingArea
4	2.0	NaN

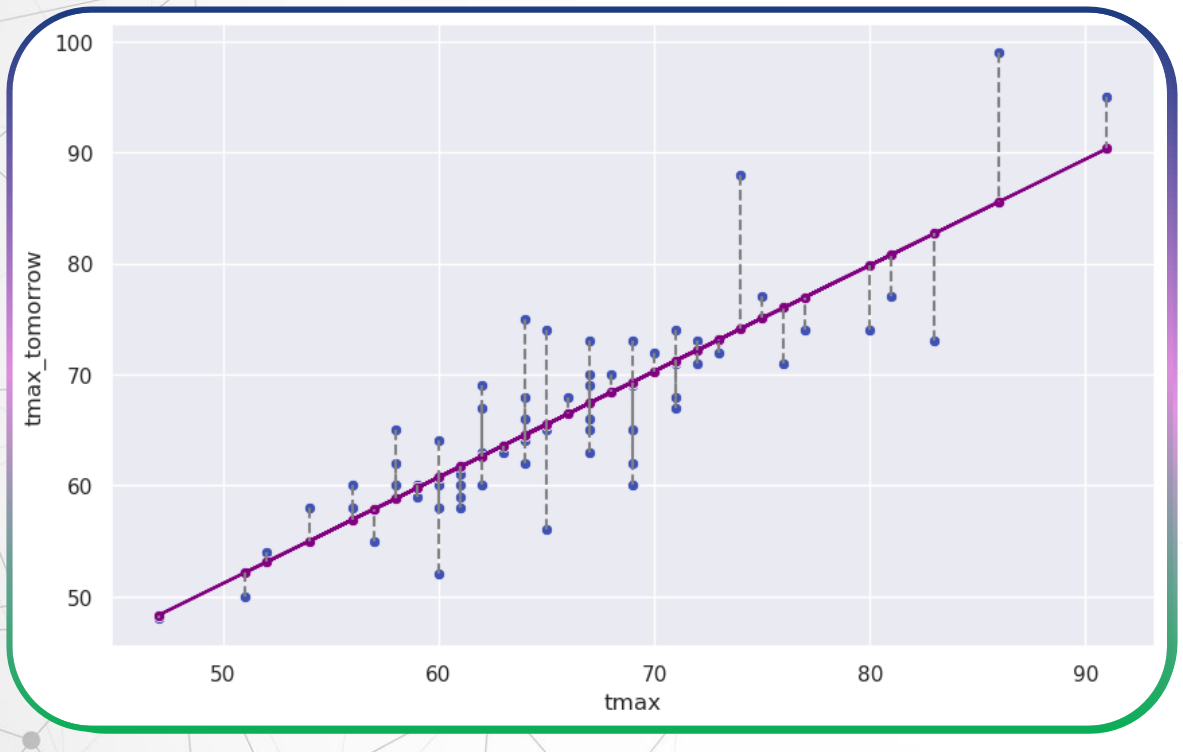
Price
124500

Supervised Regression: Linear Regression

	tmax	tmax_tomorrow
1970-01-01	60.0	52.0
1970-01-02	52.0	52.0
1970-01-03	52.0	53.0
1970-01-04	53.0	52.0
1970-01-05	52.0	50.0



Supervised Regression: Linear Regression



Initialize Parameters

Calculate Prediction

$$\hat{y} = mx + c$$

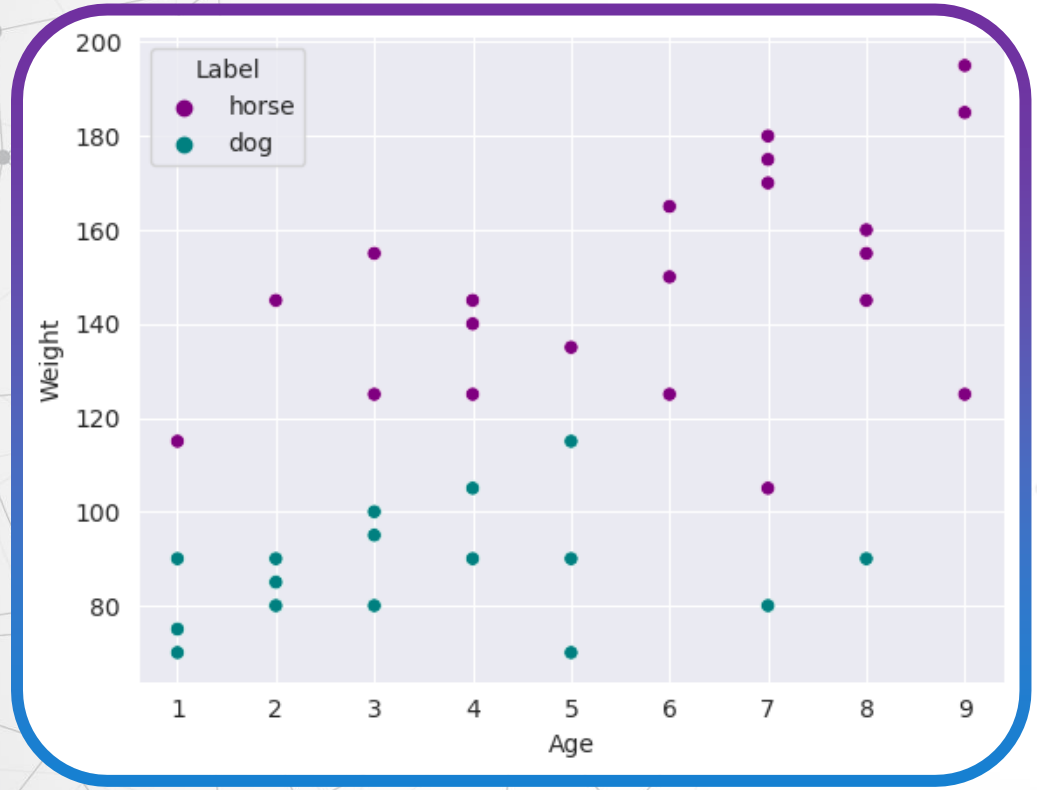
Estimate Error

$$J(m, c) = (\hat{y} - y_{actual})^2$$

Update Parameters

$$m = m - \frac{\partial J}{\partial m} \qquad c = c - \frac{\partial J}{\partial c}$$

Dog vs Horse



Age	Weight
3	150
4	120
2	90
5	200
2	85

Label
horse
horse
dog
horse
dog

Age	Weight
3	120

?

Regression vs Classification

	tmax
1970-01-01	60.0
1970-01-02	52.0
1970-01-03	52.0
1970-01-04	53.0
1970-01-05	52.0

tmax_tomorrow
52.0
52.0
53.0
52.0
50.0

Age	Weight
3	150
4	120
2	90
5	200
2	85

Label
horse
horse
dog
horse
dog

```
df['New_Label'] = df.Label.map({'horse':1, 'dog': 0})
```

Age	Weight
3	150
4	120
2	90
5	200
2	85

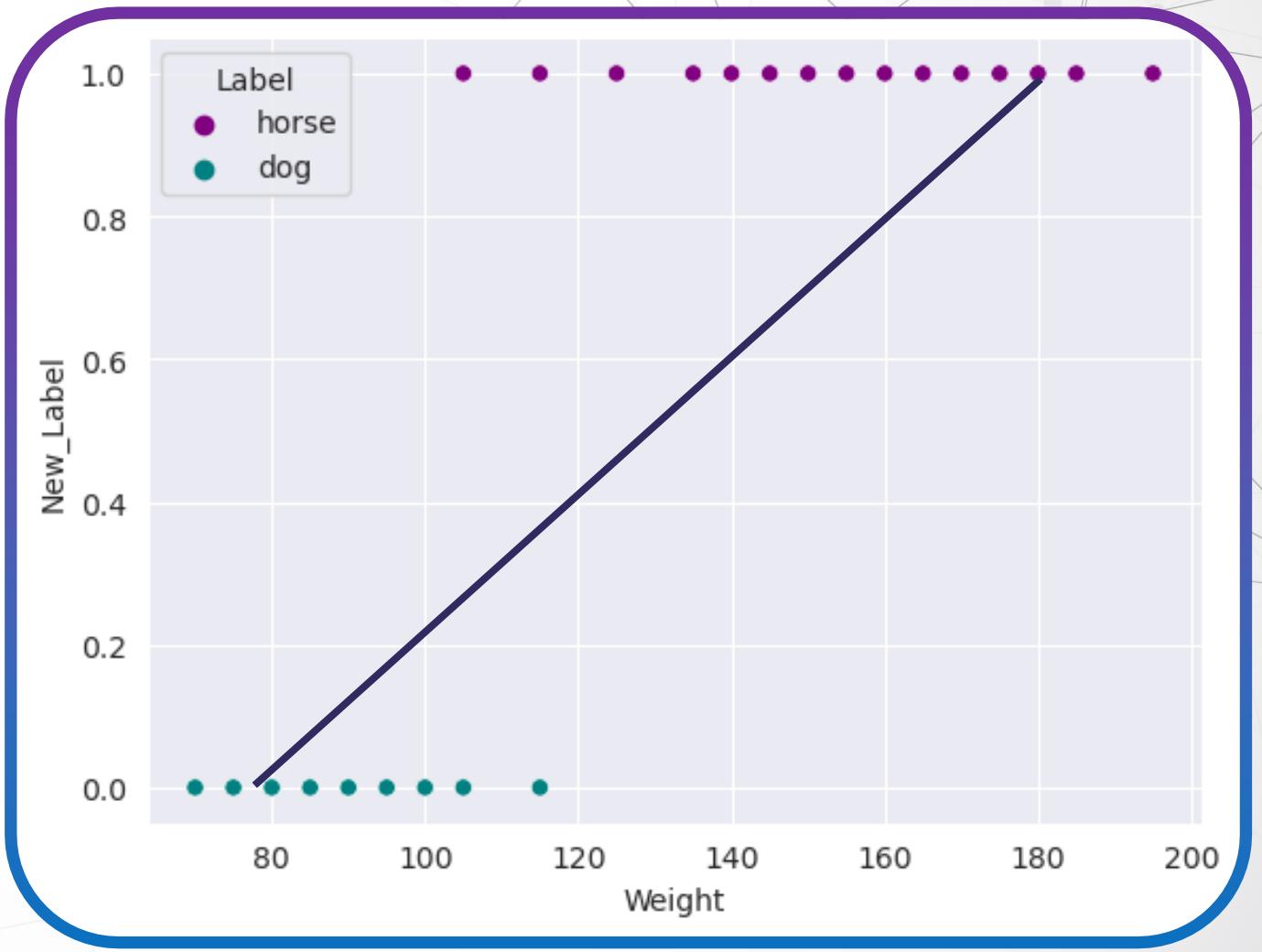
Label
horse
horse
dog
horse
dog

New_Label
1
1
0
1
0

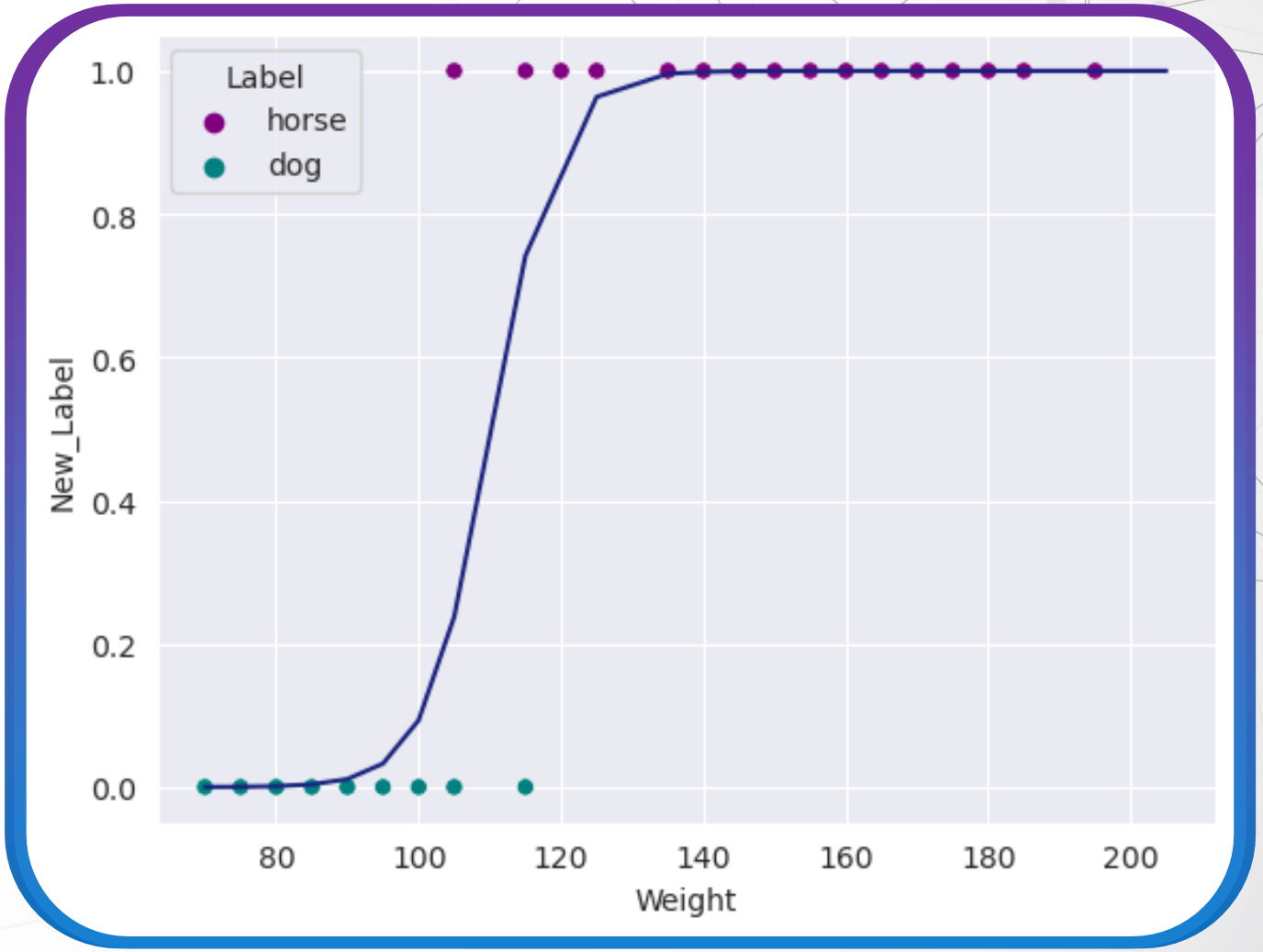
Horse -> 1

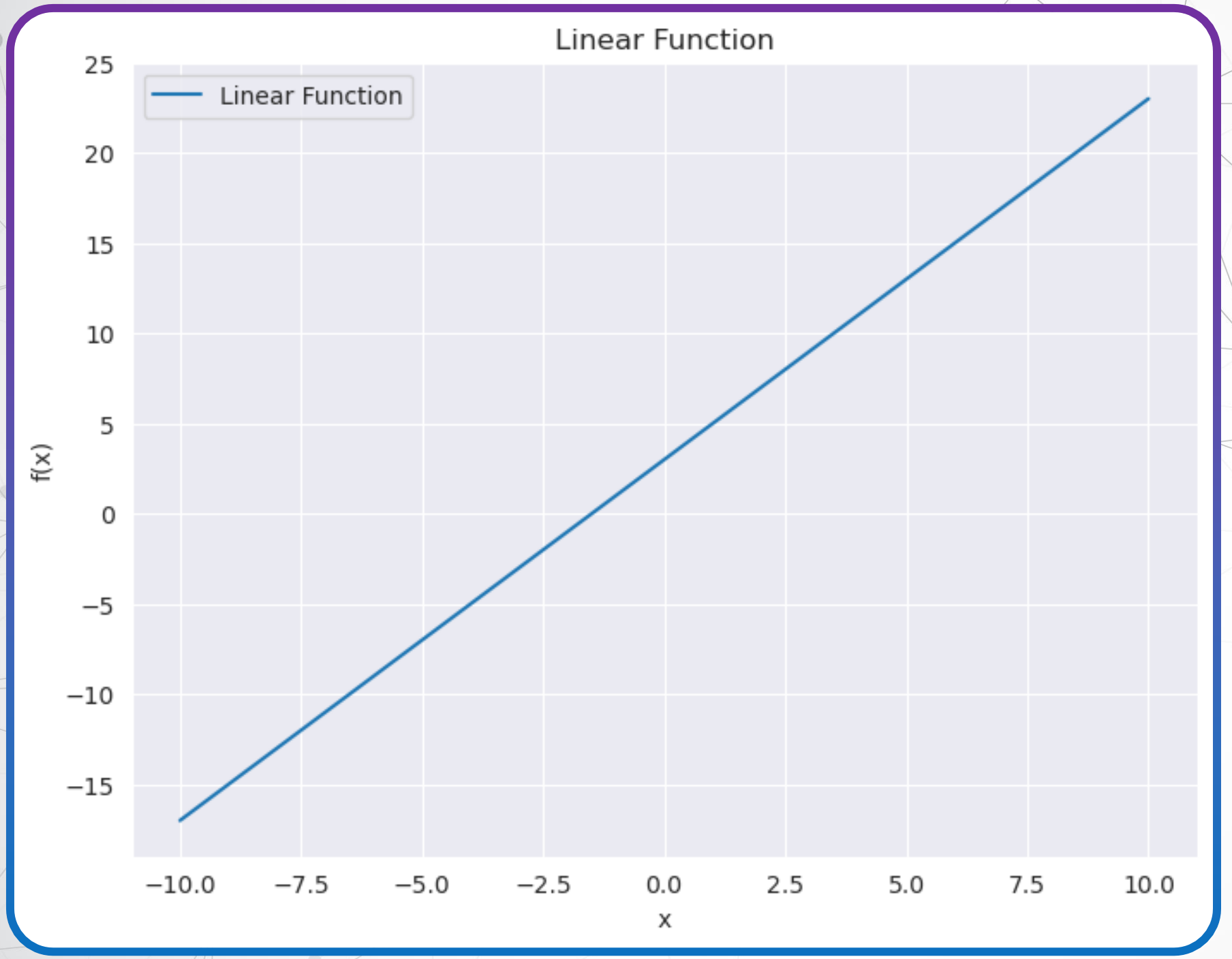
Dog -> 0

Weight	New_Label
150	1
120	1
90	0
200	1
85	0

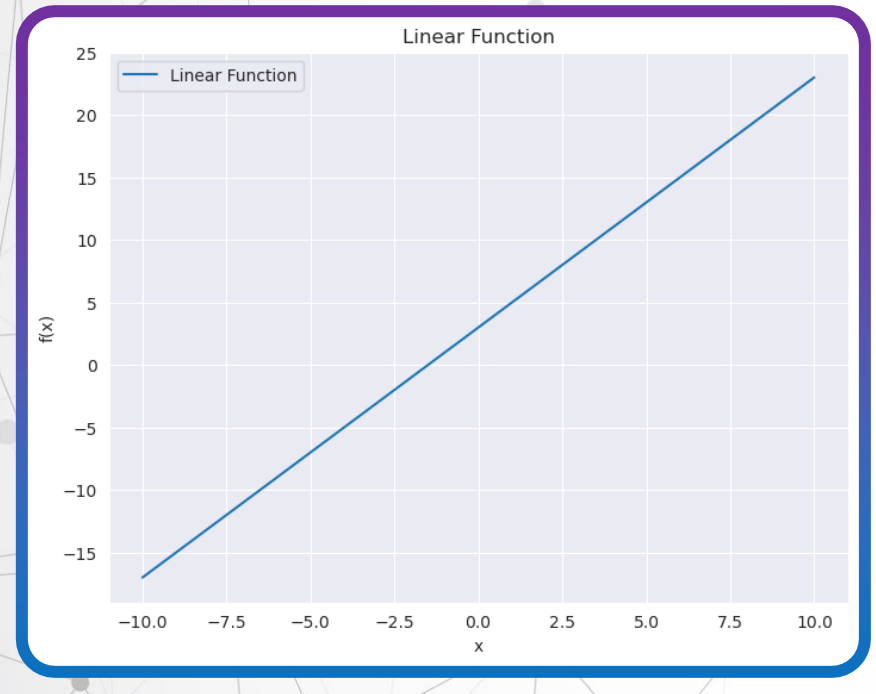


Weight	New_Label
150	1
120	1
90	0
200	1
85	0

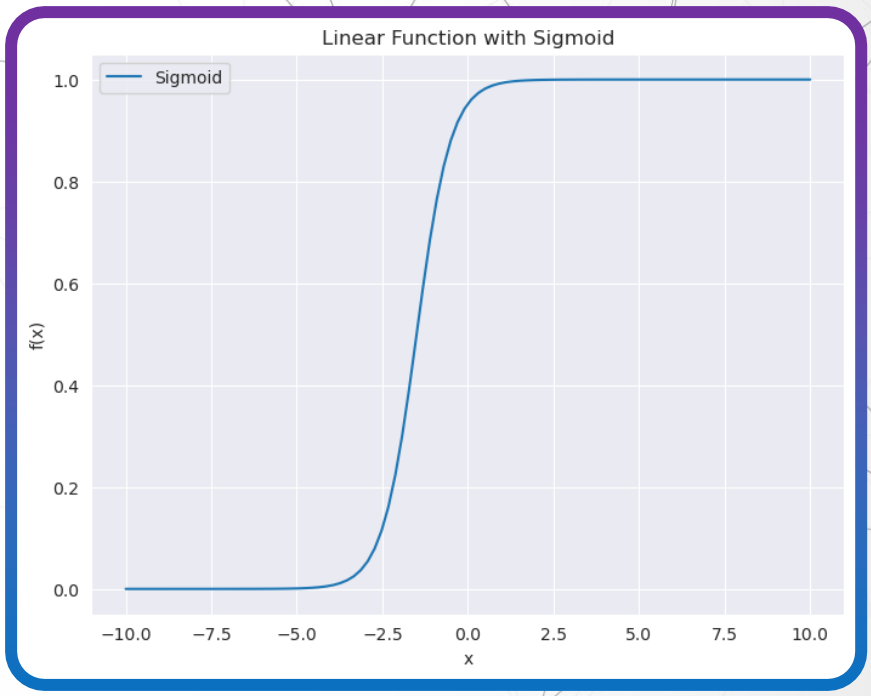



Linear Function
Linear Function
25
20
15
10
5
0
−5
−10
−15
f(x)
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
x

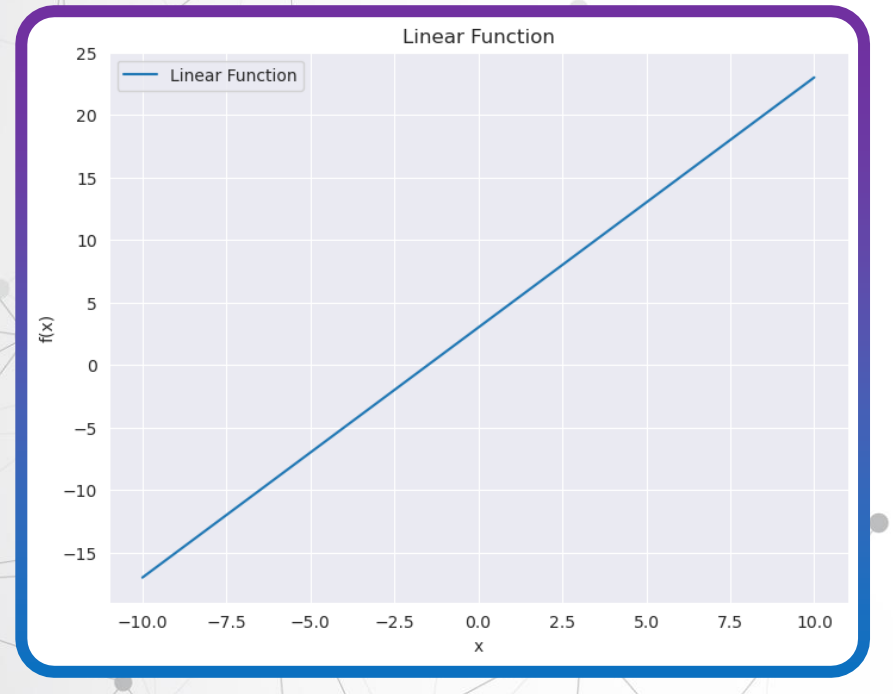
Activation Function



$$f(x) = \frac{1}{1 + e^{-x}}$$

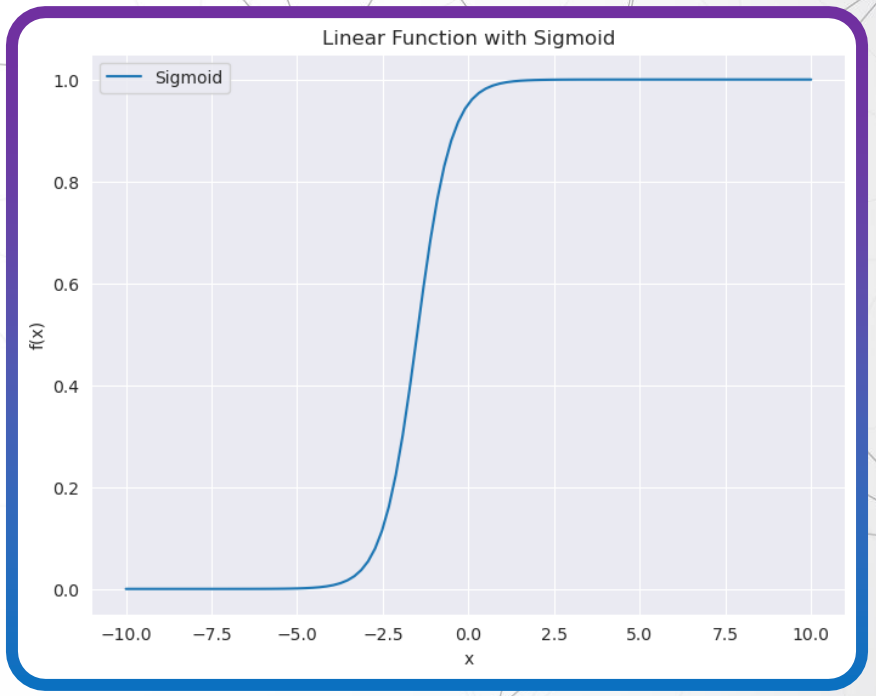


Activation Function

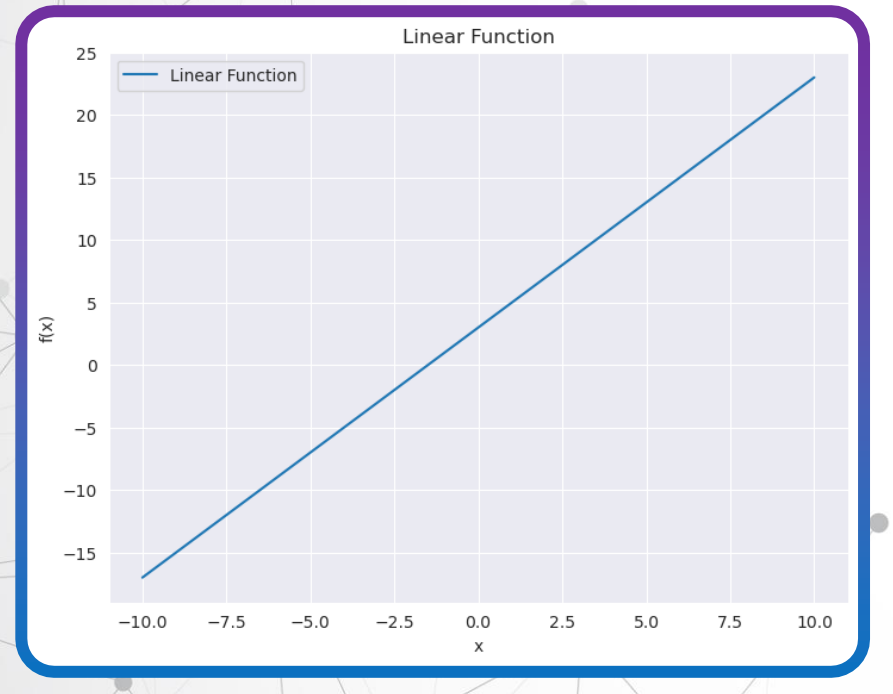


$$f(x) = \frac{1}{1 + e^{-x}}$$

$$x \to \infty, f(x) \to 1$$
$$x \to -\infty, f(x) \to 0$$

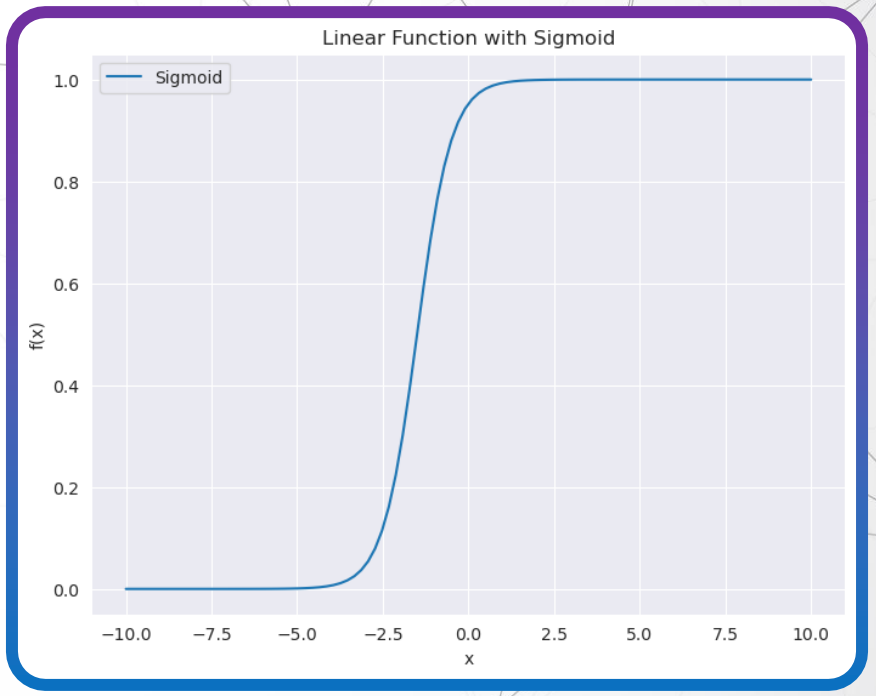


Activation Function: Sigmoid



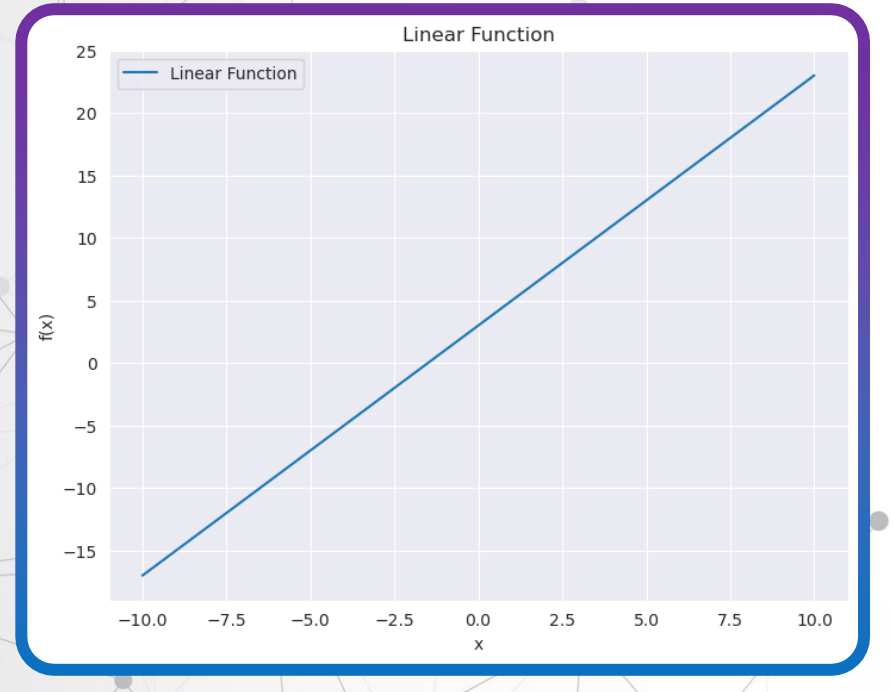
$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$x \to \infty, f(x) \to 1$$
$$x \to -\infty, f(x) \to 0$$



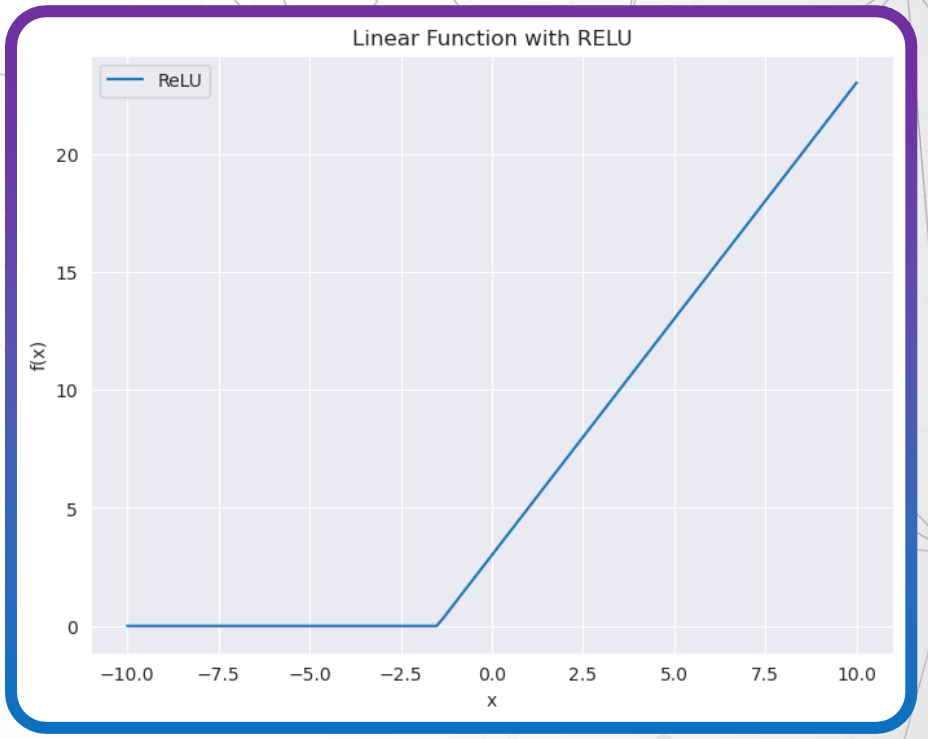
$2x + 3$

Other Activation Functions: ReLU

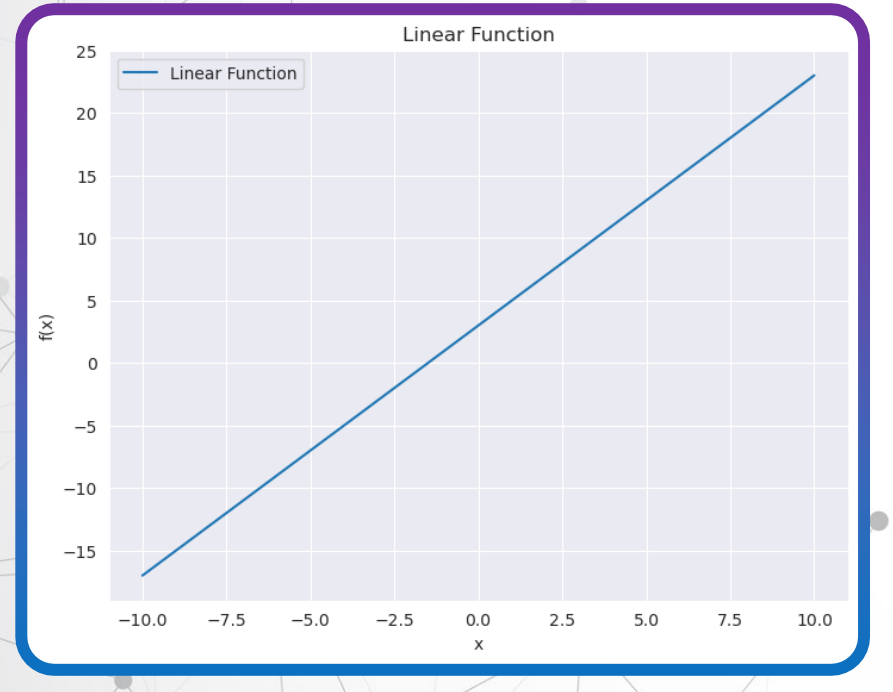


$$f(x) = max(0, x)$$

$$x < 0, f(x) = 0$$
$$x \geq 0, f(x) = x$$

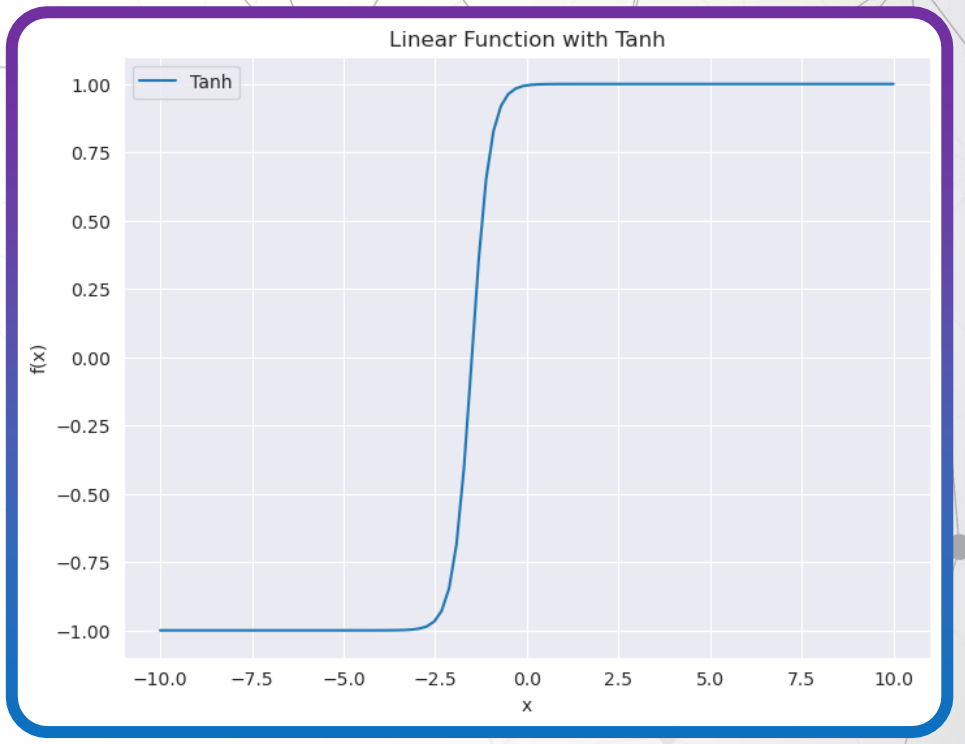


Other Activation Functions: Tanh

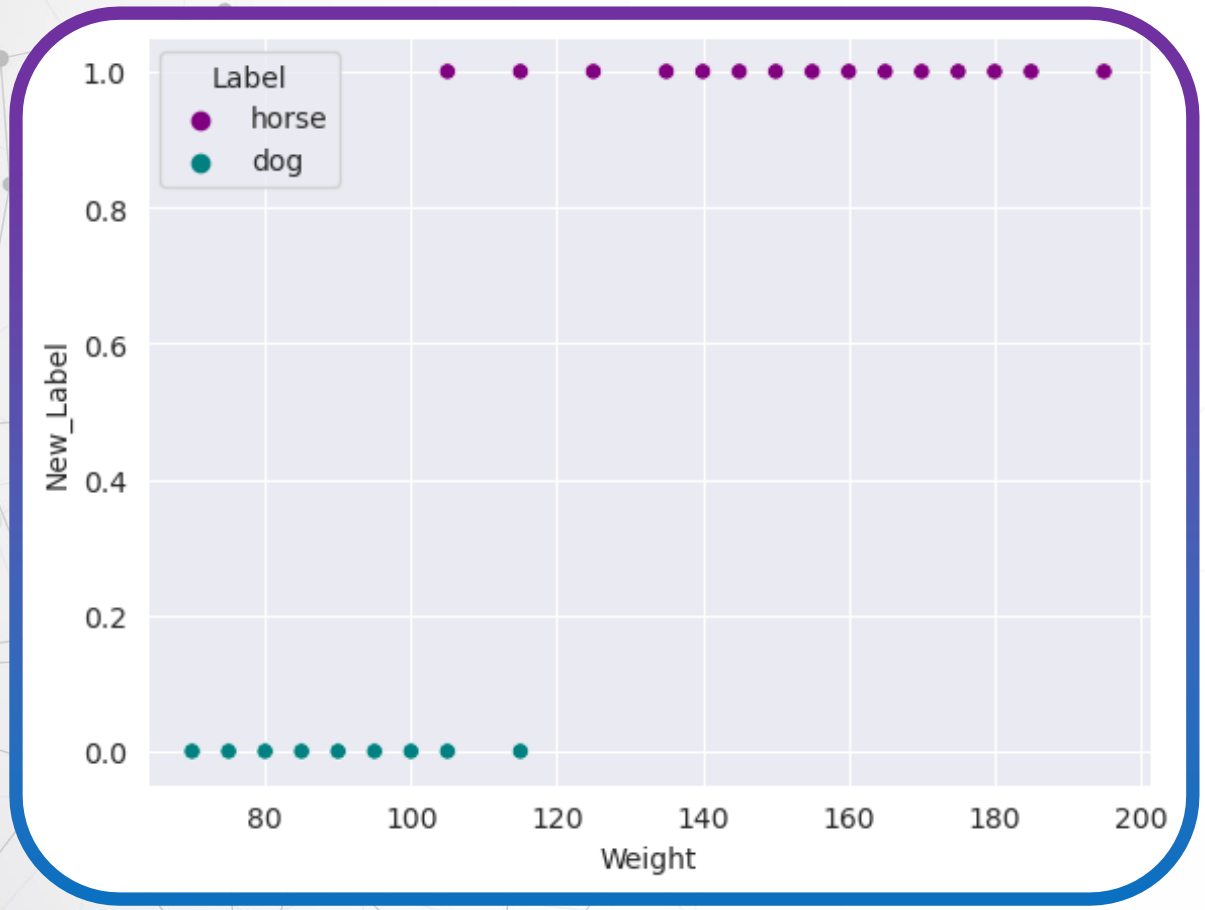


$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

$$x \to \infty, f(x) \to 1$$
$$x \to -\infty, f(x) \to -1$$



Supervised Regression



Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(mx + c)$$

Estimate Error

$$J(m, c) = (\hat{y} - y_{actual})^2$$

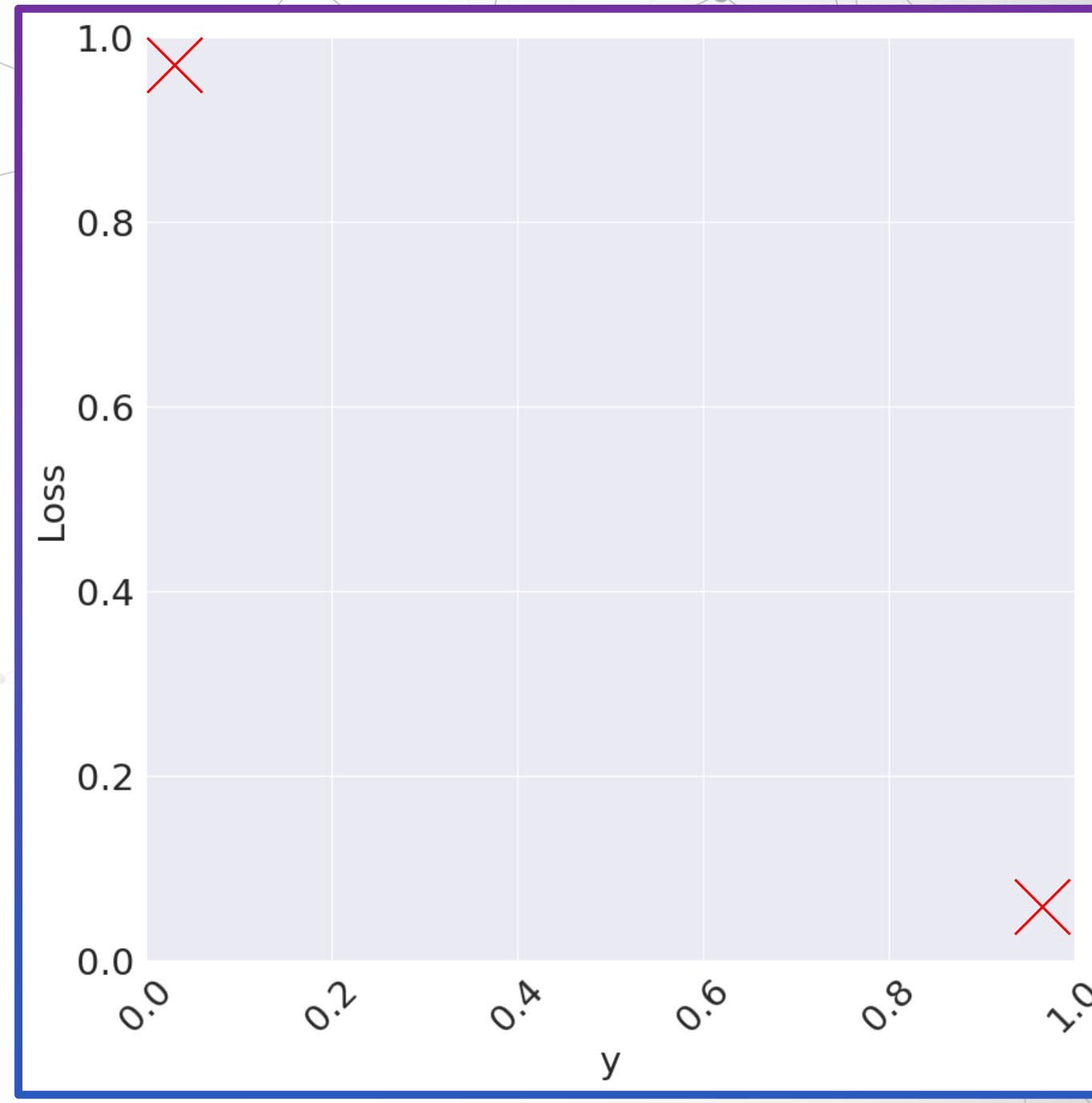
Update Parameters

$$m = m - \frac{\partial J}{\partial m} \qquad c = c - \frac{\partial J}{\partial c}$$

When $\boldsymbol{y = 1}$

if $\boldsymbol{\hat{y} = 1}$, loss should be 0

if $\boldsymbol{\hat{y} \to 0}$, loss should be high value

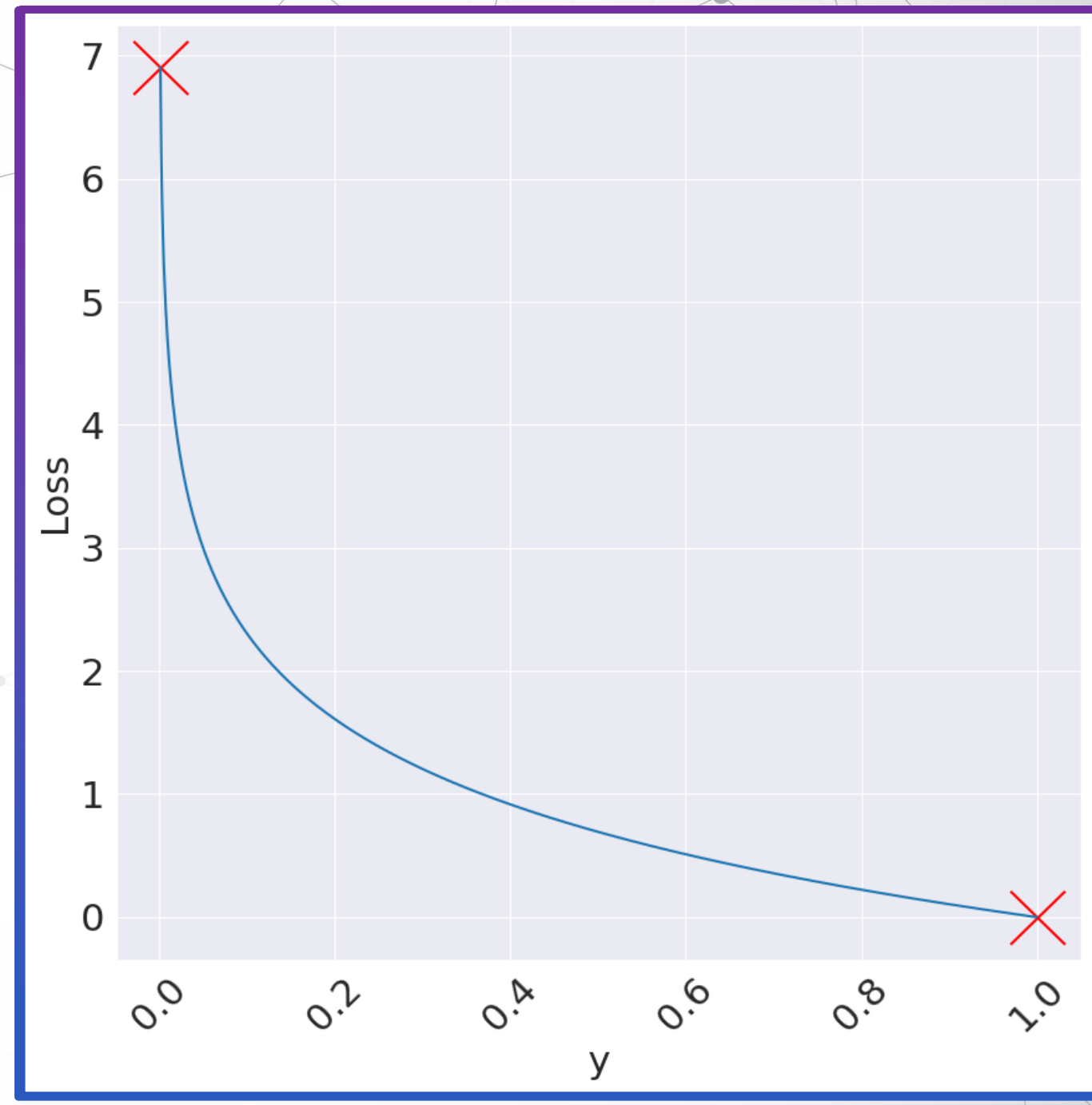


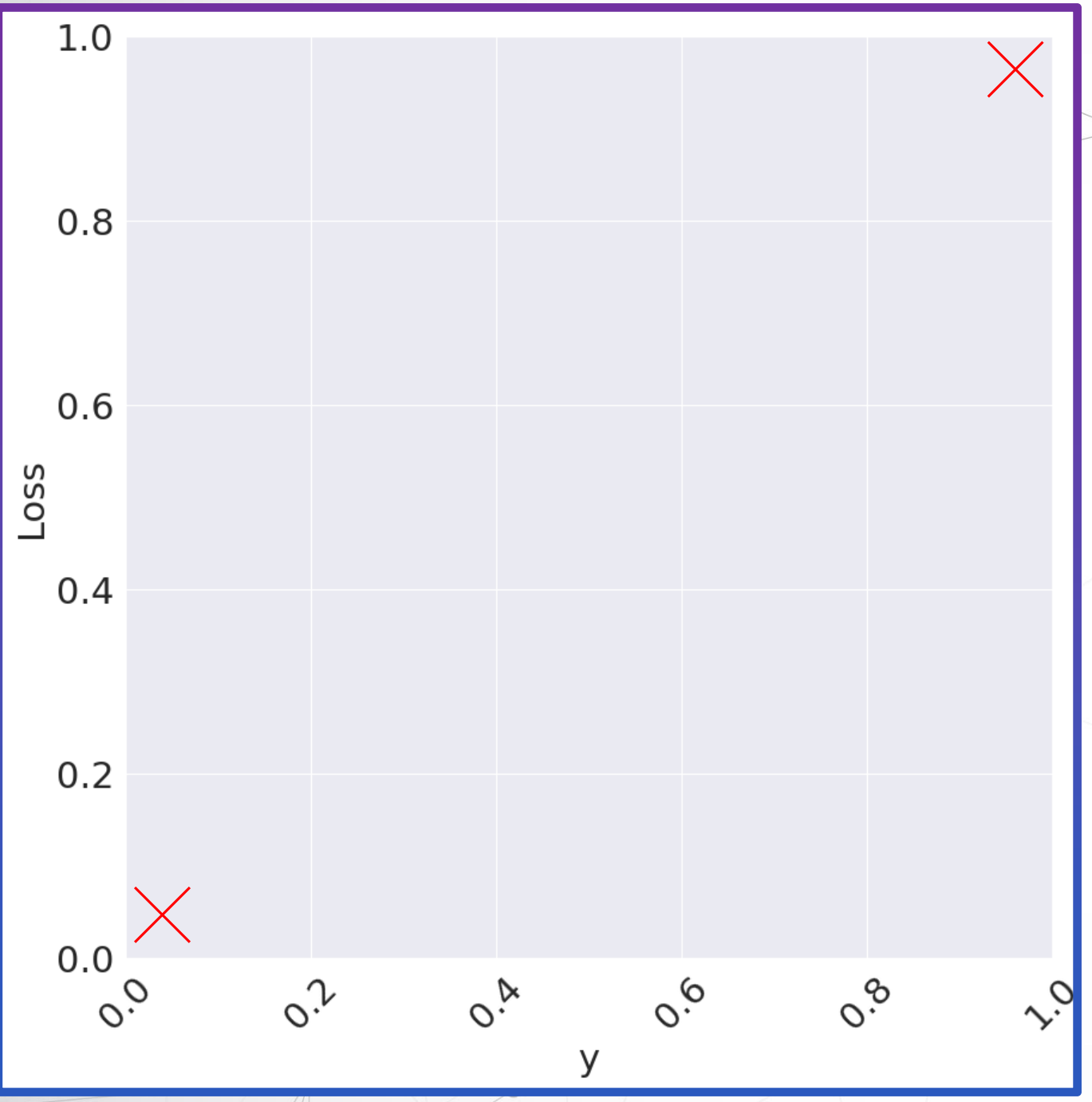
When $y = 1$

if $\hat{y} = 1$, loss should be 0

if $\hat{y} \to 0$, loss should be high value

$$Loss = -\log(\hat{y})$$



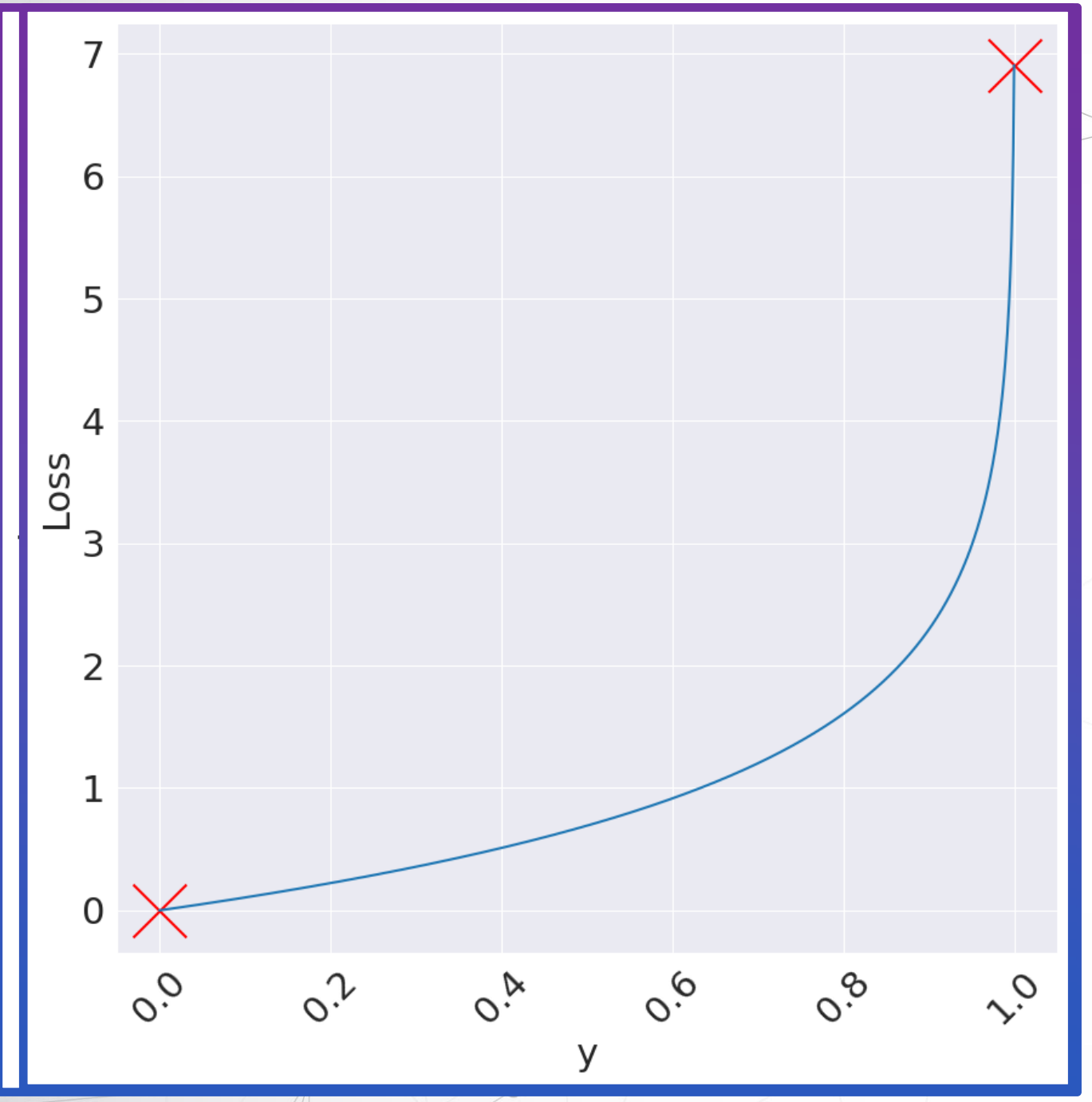


When $\boldsymbol{y} = \mathbf{0}$

if $\hat{\boldsymbol{y}} = \mathbf{0}$, loss should be 0

if $\hat{\boldsymbol{y}} \to \mathbf{1}$, loss should be high value

$$Loss = -\log(1 - \hat{\boldsymbol{y}})$$



When $\boldsymbol{y} = \mathbf{0}$

if $\hat{\boldsymbol{y}} = \mathbf{0}$, loss should be 0

if $\hat{\boldsymbol{y}} \to \mathbf{1}$, loss should be high value

$$Loss = -\log(1 - \hat{\boldsymbol{y}})$$

When $\boldsymbol{y = 1}$

if $\hat{\boldsymbol{y}} = \mathbf{1}$, loss should be 0

if $\hat{\boldsymbol{y}} \to \mathbf{0}$, loss should be high value

When $\boldsymbol{y = 0}$

if $\hat{\boldsymbol{y}} = \mathbf{0}$, loss should be 0

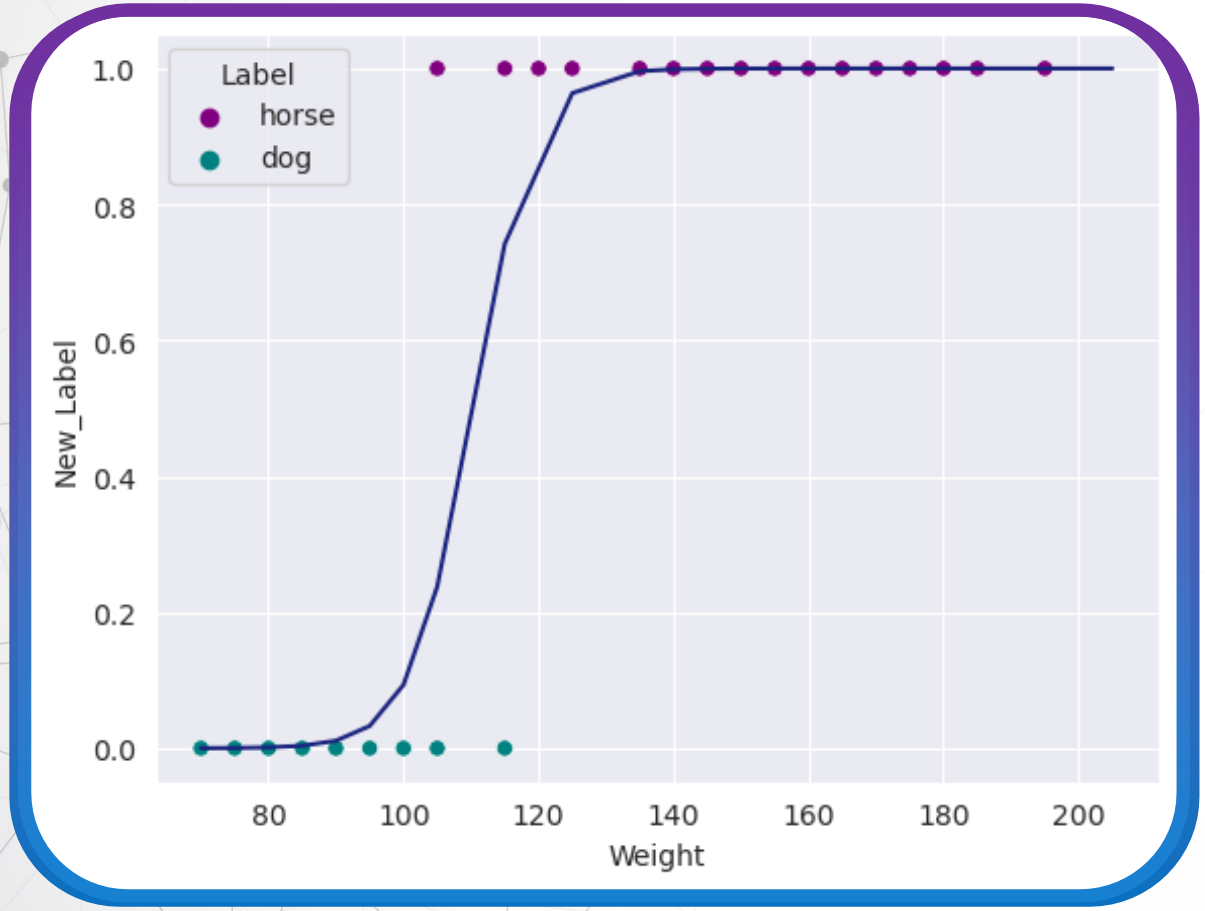
if $\hat{\boldsymbol{y}} \to \mathbf{1}$, loss should be high value

$$Loss = -\log(\hat{\boldsymbol{y}}) = -\boldsymbol{y}\log(\hat{\boldsymbol{y}})$$

$$Loss = -\log(1-\hat{\boldsymbol{y}}) = -(1-y)\log(1-\hat{\boldsymbol{y}})$$

$$Loss = -\boldsymbol{y}\log(\hat{\boldsymbol{y}}) - (1-y)\log(1-\hat{\boldsymbol{y}})$$

Supervised Regression: Logistic Regression



Initialize Parameters

Calculate Prediction

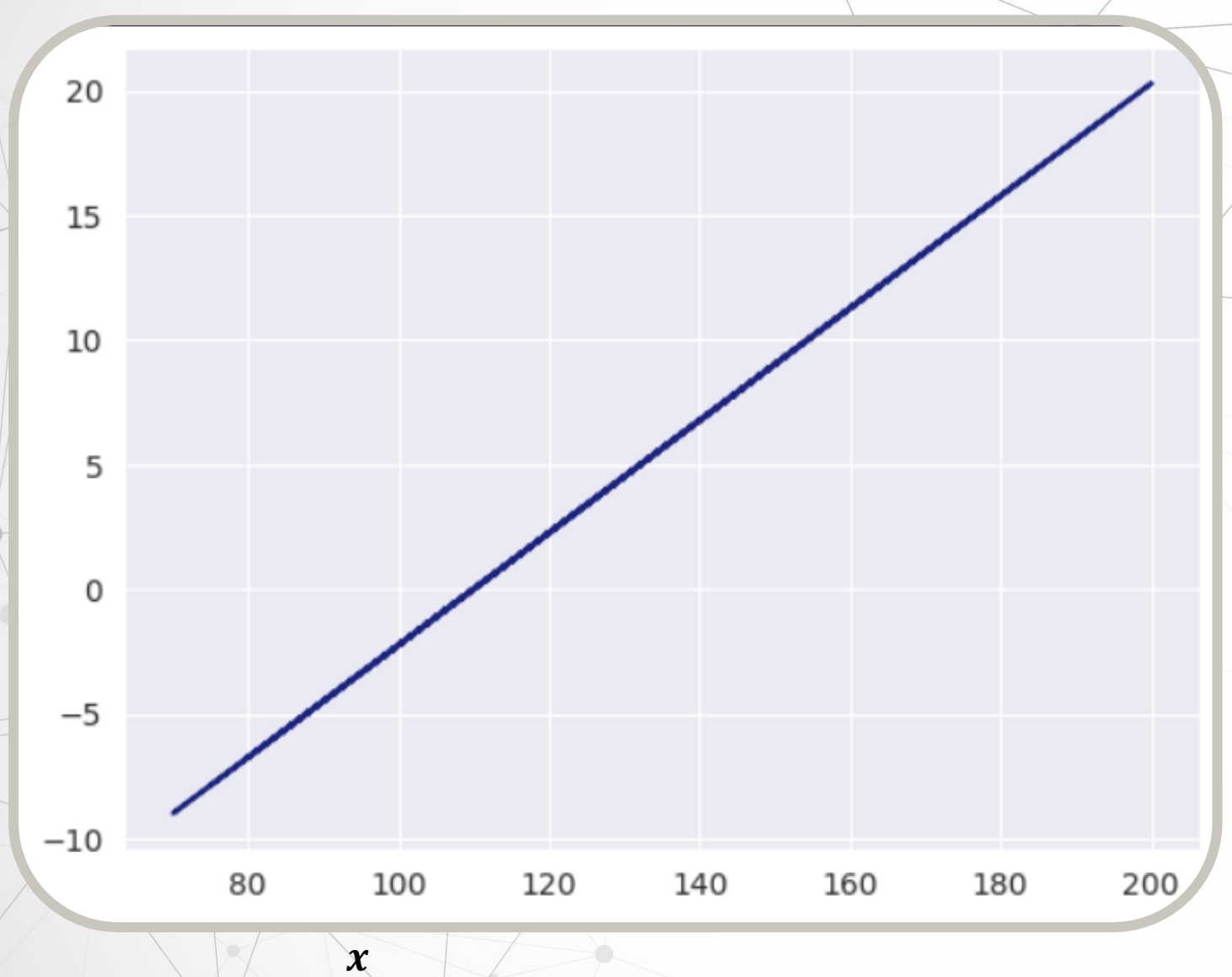
$$\hat{y} = \sigma(mx + c)$$

Estimate Error

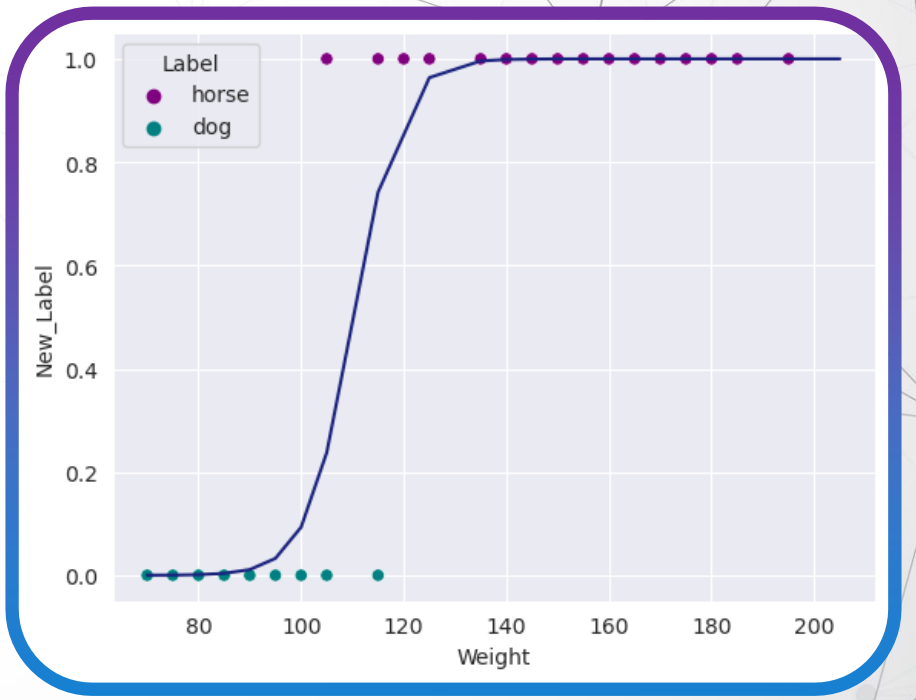
$$J(m, c) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

$$m = m - \frac{\partial J}{\partial m} \qquad c = c - \frac{\partial J}{\partial c}$$



$$\sigma(x) = \frac{1}{1 + e^{-x}}$$



tmax	tmax_tomorrow
60.0	52.0

Initialize Parameters

Calculate Prediction

$$\hat{y} = mx + c$$

Estimate Error

$$J(m, c) = (\hat{y} - y_{actual})^2$$

Update Parameters

$$m = m - \frac{\partial J}{\partial m} \quad c = c - \frac{\partial J}{\partial c}$$

Weight	New_Label
150	1

Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(mx + c)$$

Estimate Error

$$J(m, c) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

$$m = m - \frac{\partial J}{\partial m} \quad c = c - \frac{\partial J}{\partial c}$$

tmax	tmax_tomorrow
60.0	52.0

Initialize Parameters

Calculate Prediction

$$\hat{y} = \theta_0 x + \theta_1$$

Estimate Error

$$J(\theta) = (\hat{y} - y_{actual})^2$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

Weight	New_Label
150	1

Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

Estimate Error

$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

tmax
60.0

tmax_tomorrow
52.0

Initialize Parameters

Calculate Prediction

$$\hat{y} = \theta_0 x + \theta_1$$

Estimate Error

$$J(\theta) = (\hat{y} - y_{actual})^2$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

Age
3

New_Label
1

Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

Estimate Error

$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

tmax	tmax_tomorrow
60.0	52.0

Initialize Parameters

Calculate Prediction

$$\hat{y} = \theta_0 x + \theta_1$$

Estimate Error

$$J(\theta) = (\hat{y} - y_{actual})^2$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

Weight	New_Label
150	1

Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

Estimate Error

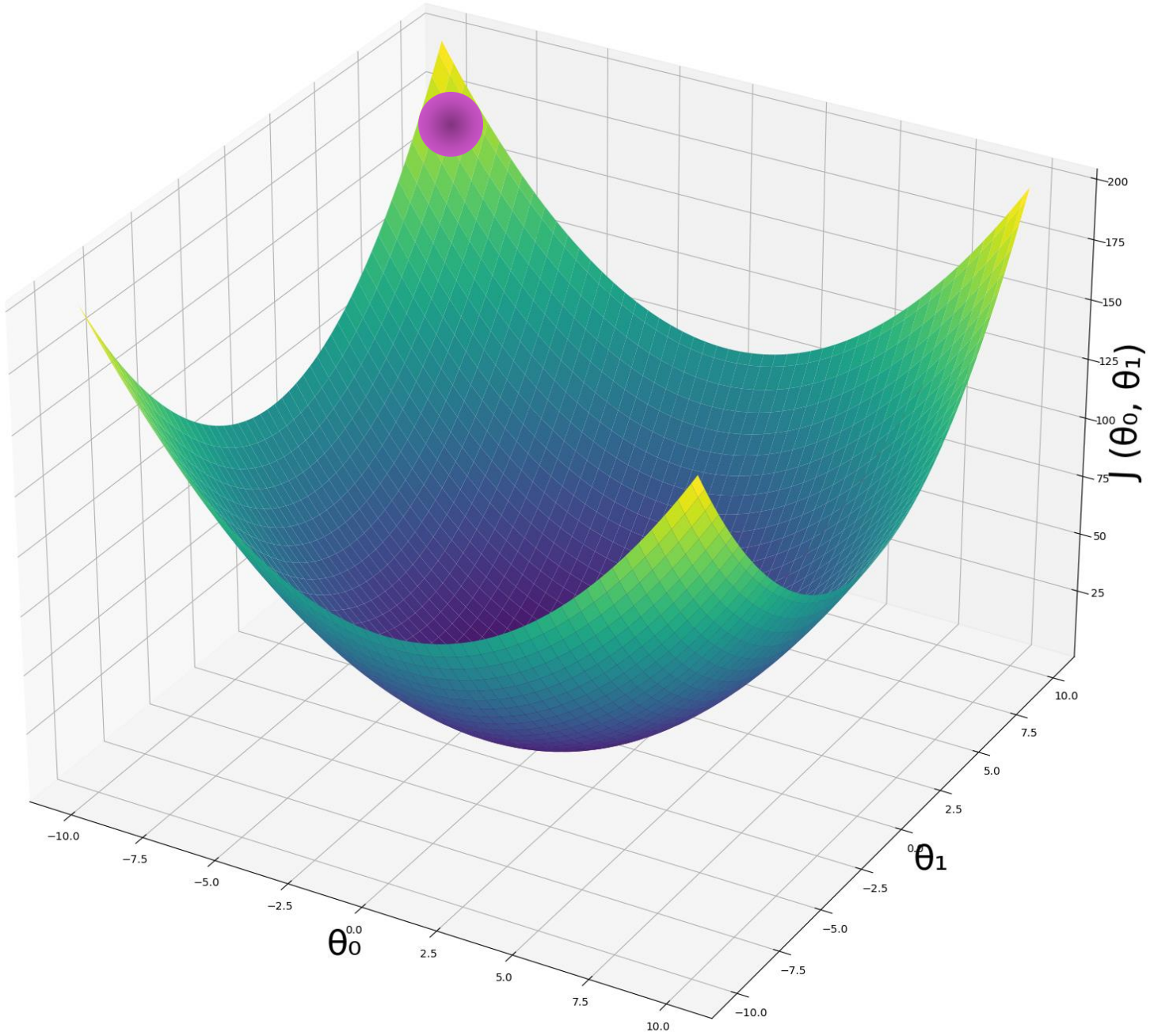
$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

Steps to modeling Machine Learning Problems

- Data: Pick your features
- Model:
  - Pick your Model
  - Pick your Loss
- Let the optimizer handle the rest



Initialize Parameters

Calculate Prediction

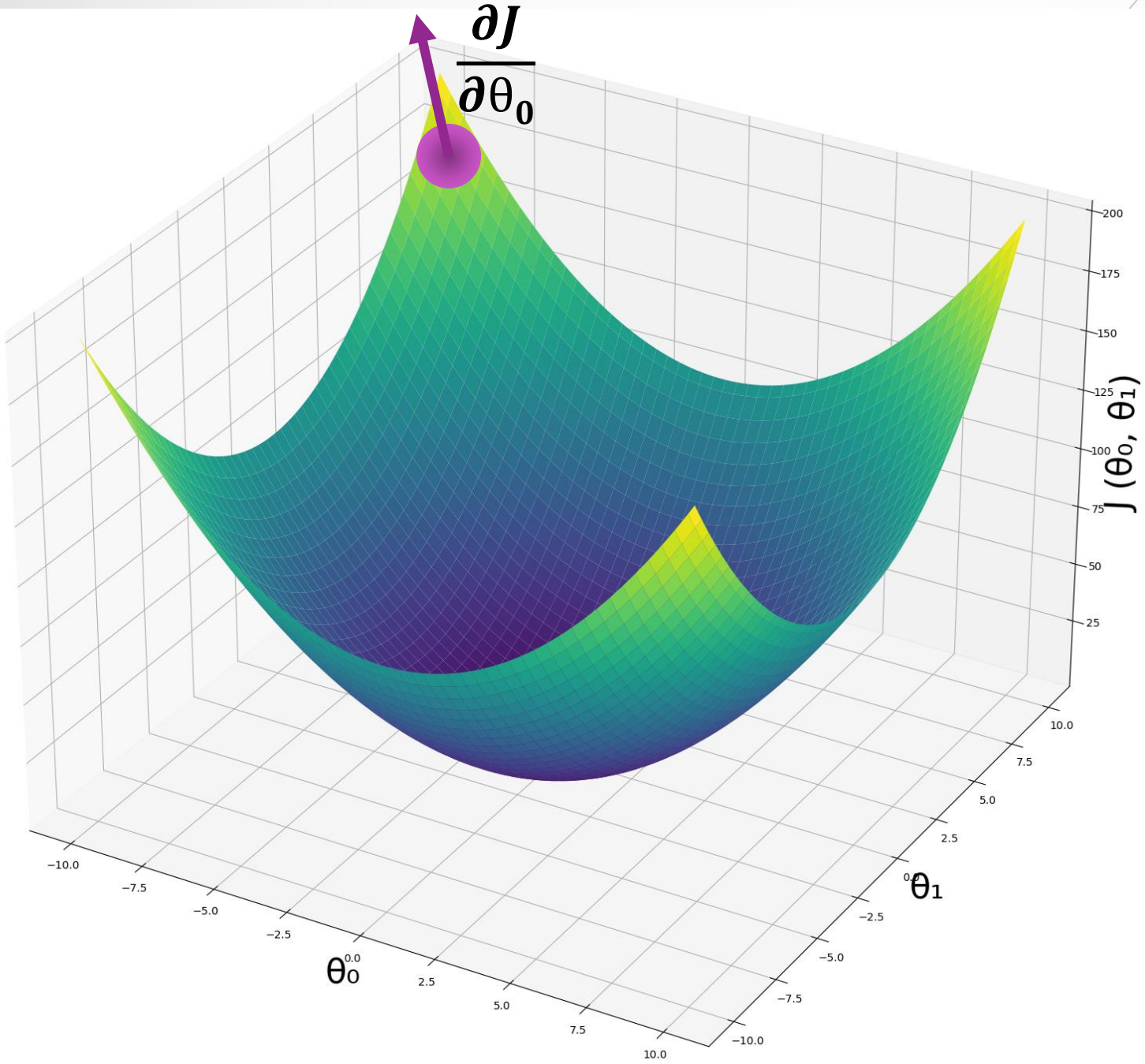
$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

Estimate Error

$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$



Initialize Parameters

Calculate Prediction

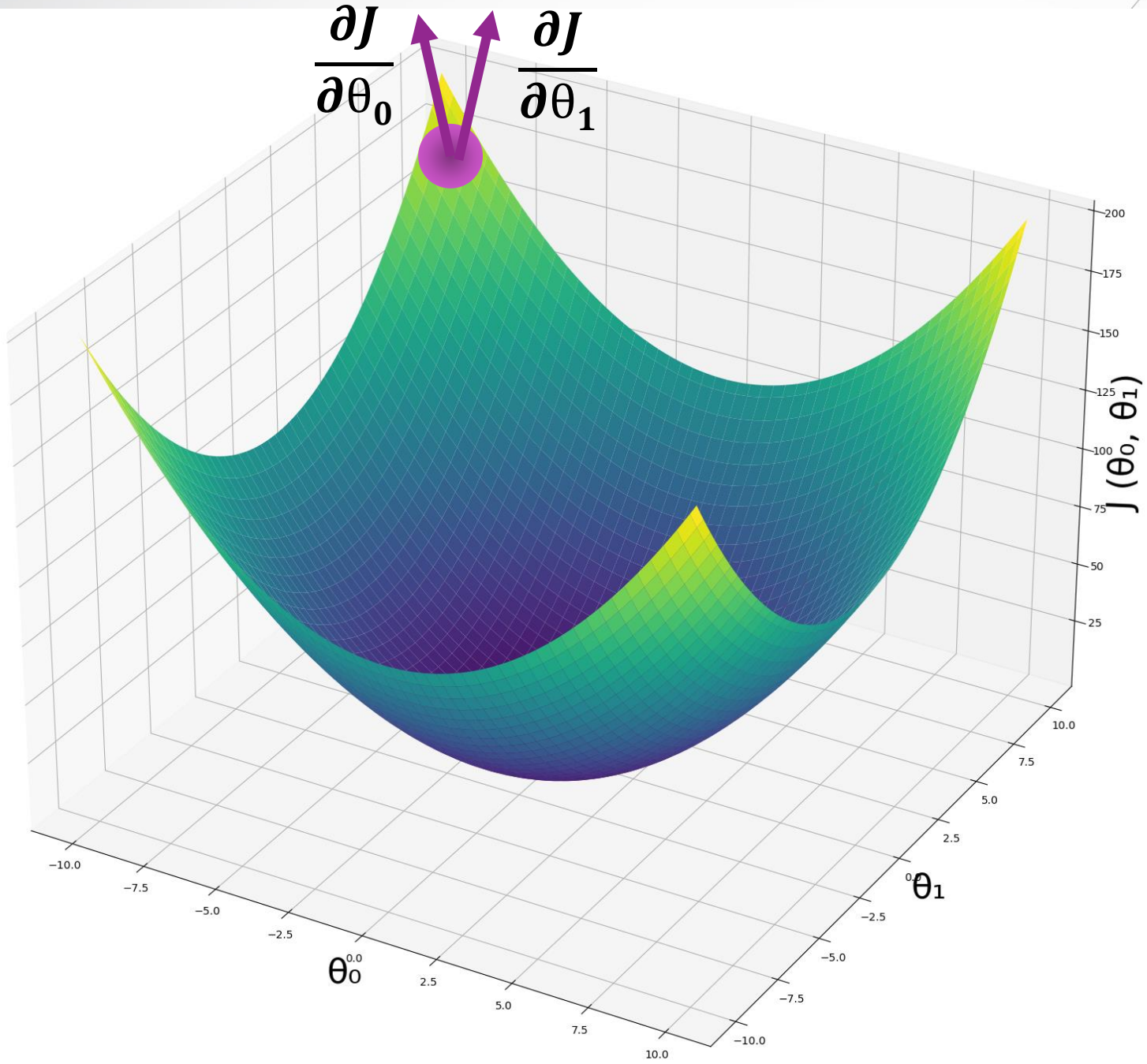
$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

Estimate Error

$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

Calculate $\frac{\partial J}{\partial \theta_0}$



Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

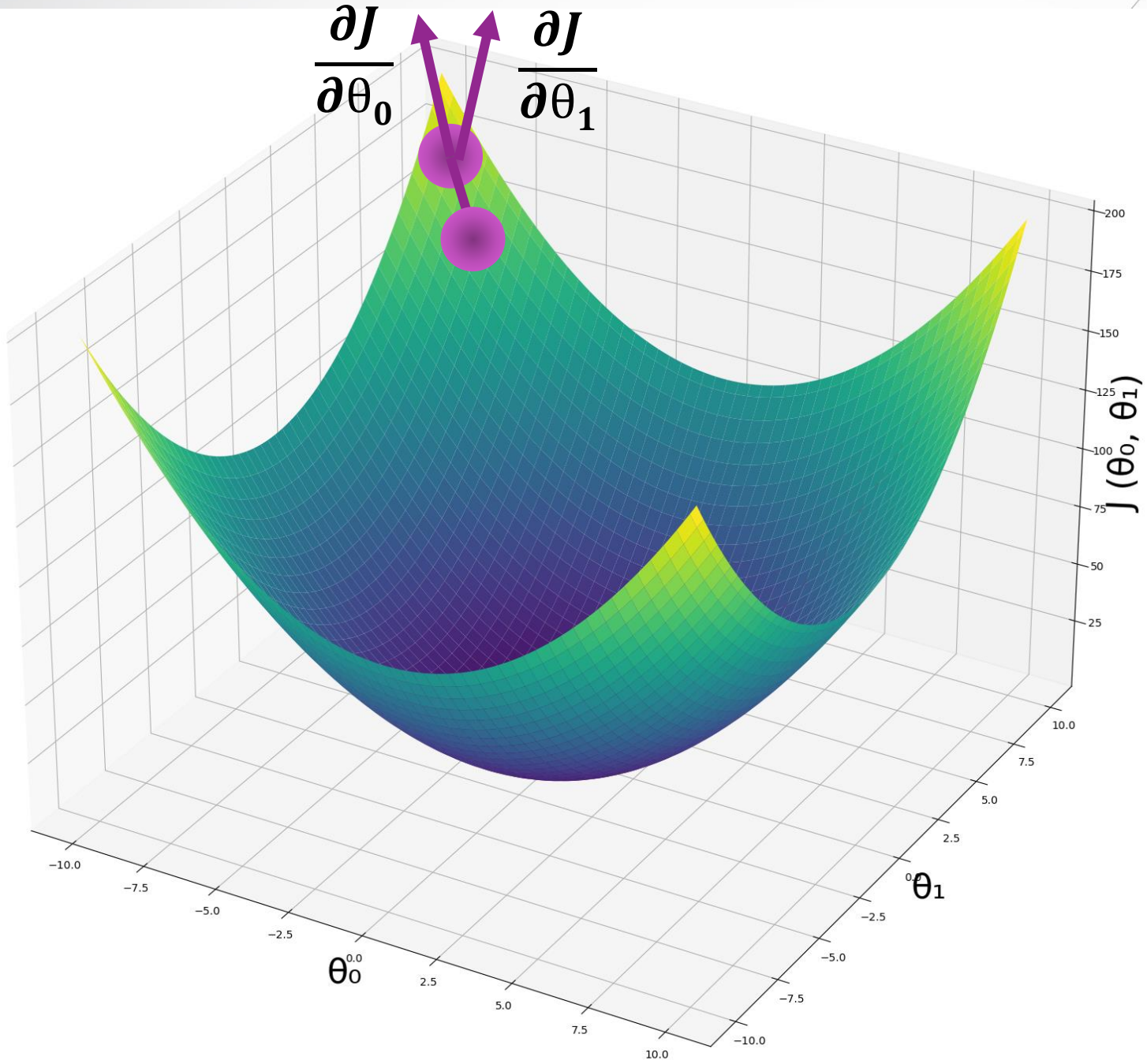
Estimate Error

$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

Calculate $\frac{\partial J}{\partial \theta_0}$,

Calculate $\frac{\partial J}{\partial \theta_1}$



Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

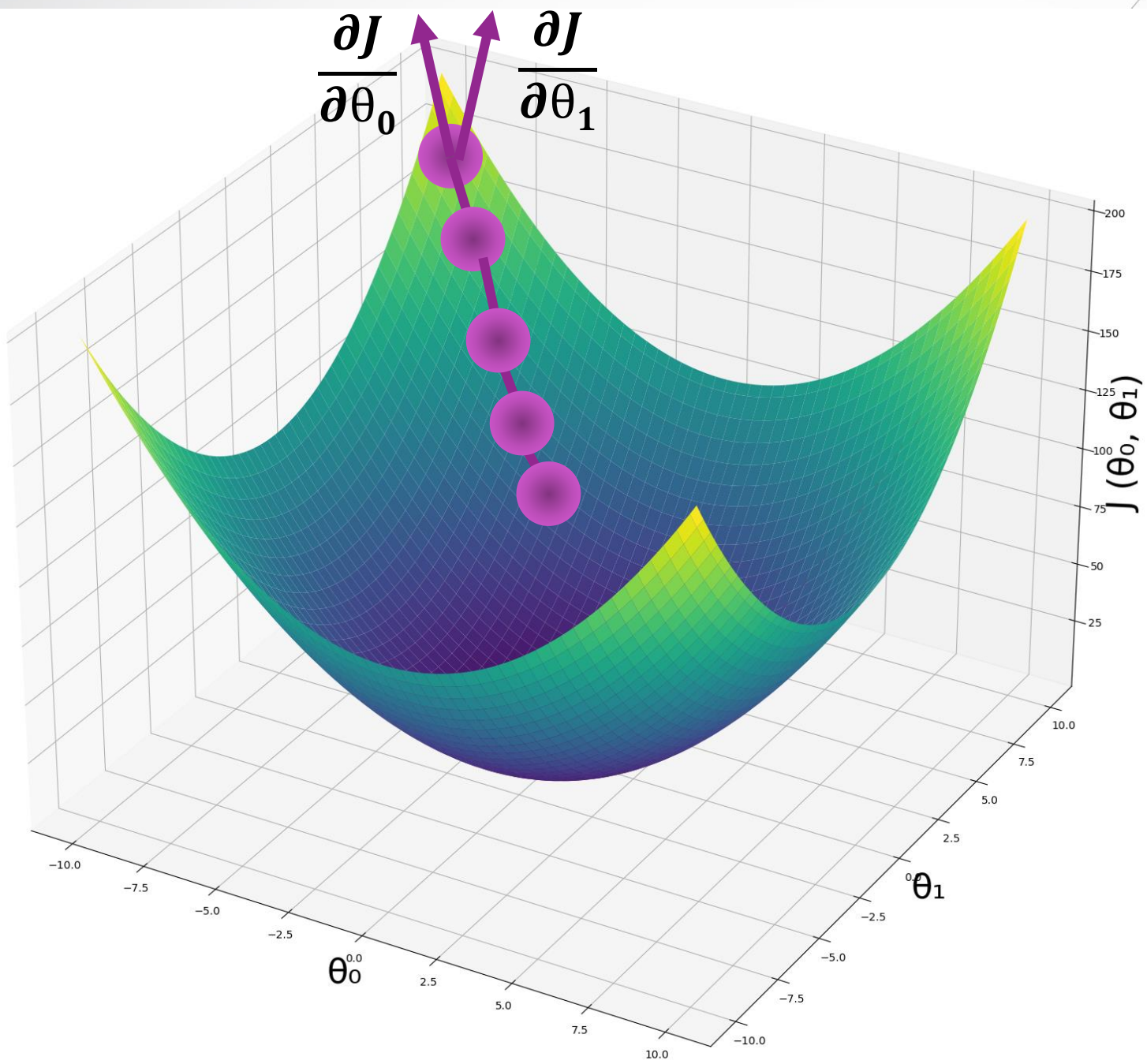
Estimate Error

$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

Calculate $\frac{\partial J}{\partial \theta_0}$, Calculate $\frac{\partial J}{\partial \theta_1}$

$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0}, \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$



Initialize Parameters

Calculate Prediction

$$\hat{y} = \sigma(\theta_0 x + \theta_1)$$

Estimate Error

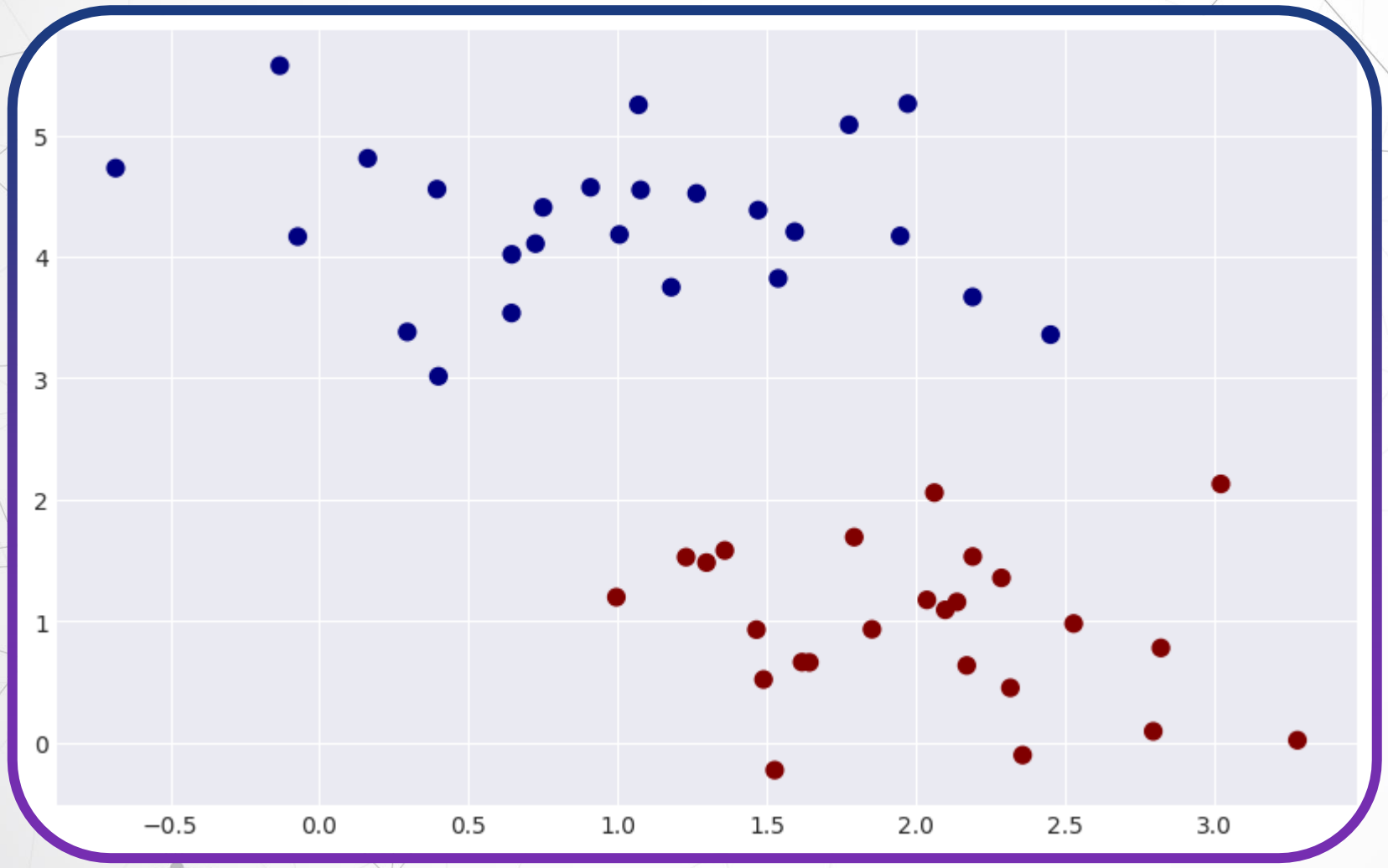
$$J(\theta) = CrossEntropy(\hat{y}, y_{actual})$$

Update Parameters

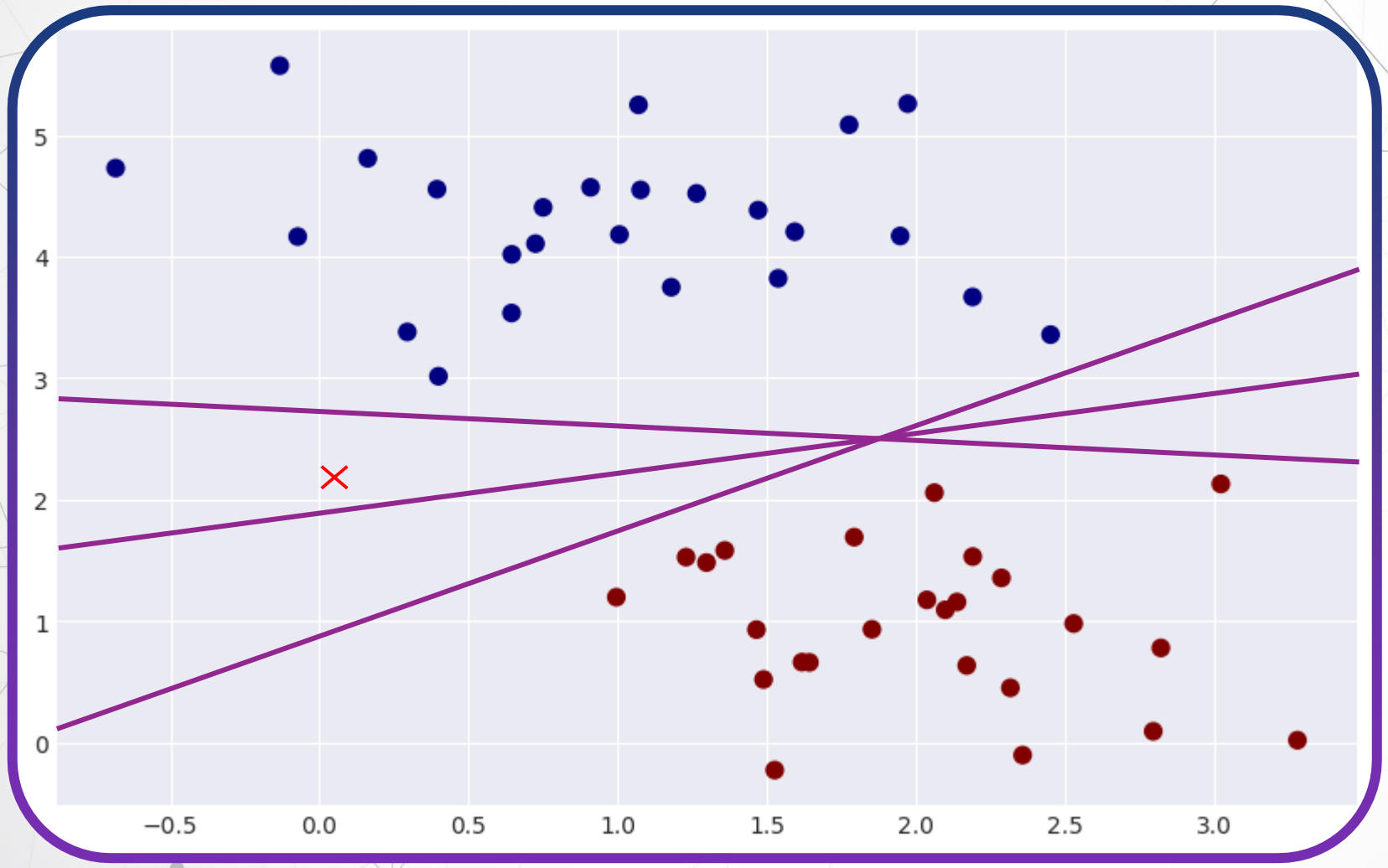
$$\theta_0 = \theta_0 - \frac{\partial J}{\partial \theta_0} \qquad \theta_1 = \theta_1 - \frac{\partial J}{\partial \theta_1}$$

Up Next:
SVM and Decision Trees

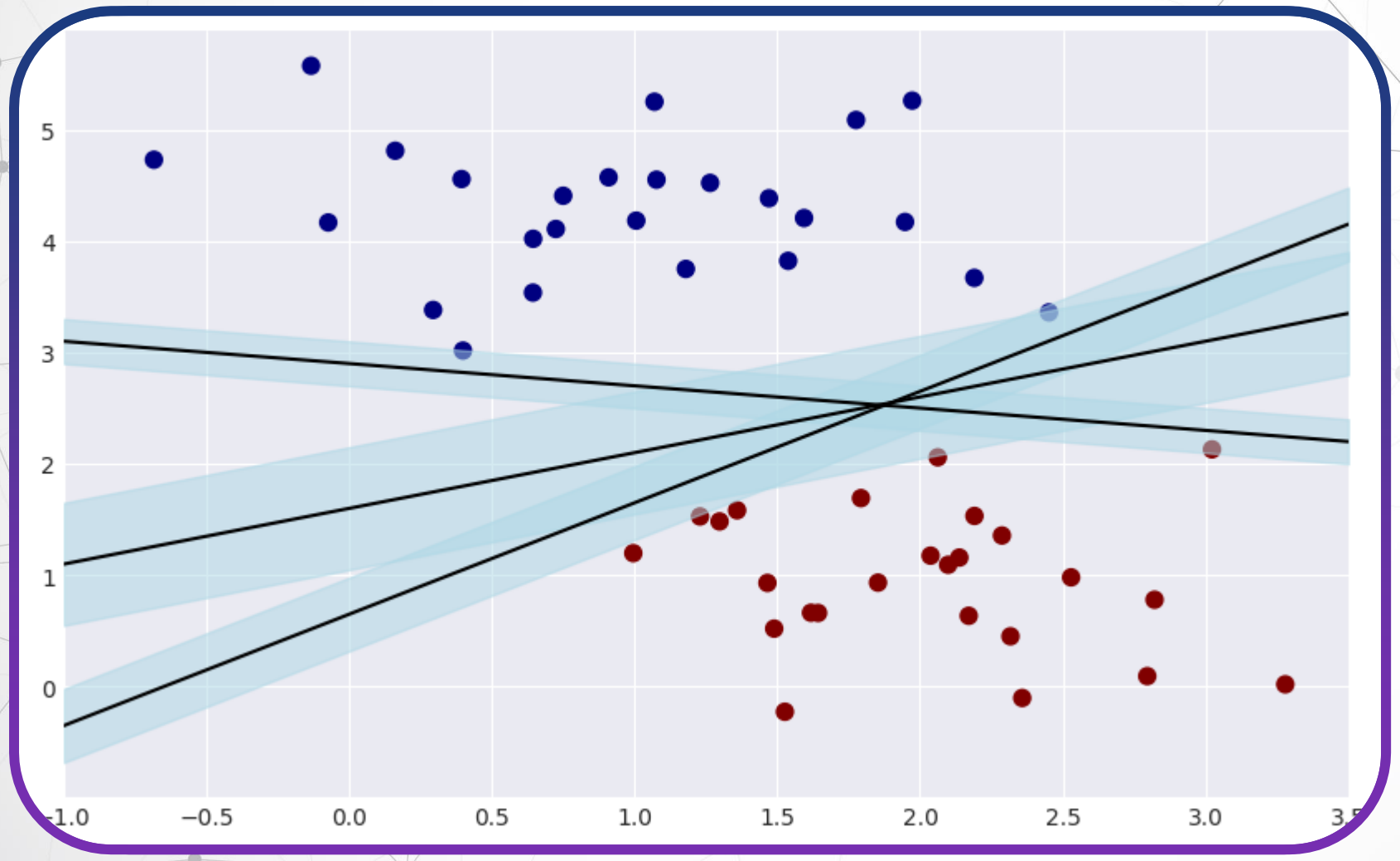
Support Vector Machines



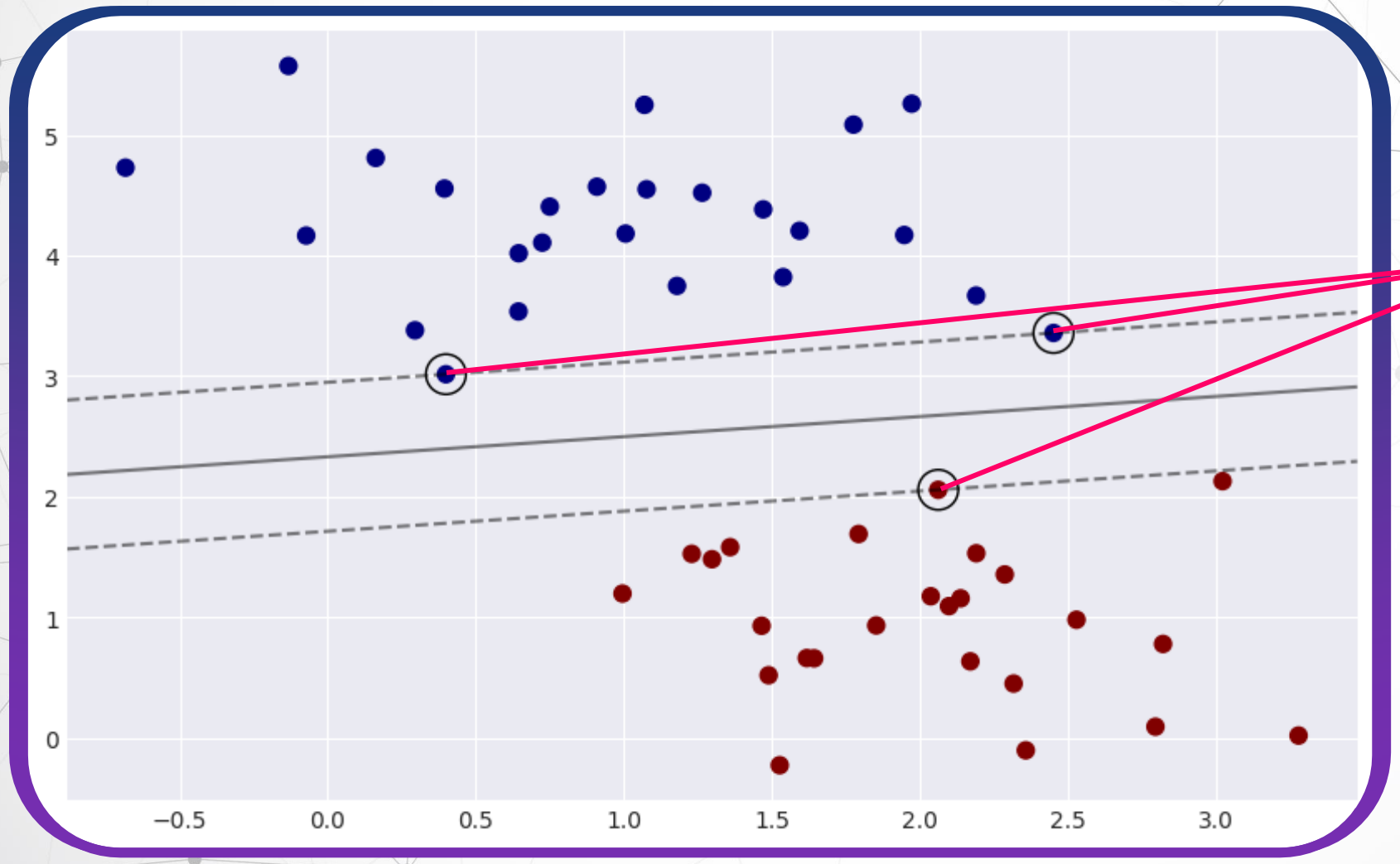
Support Vector Machines



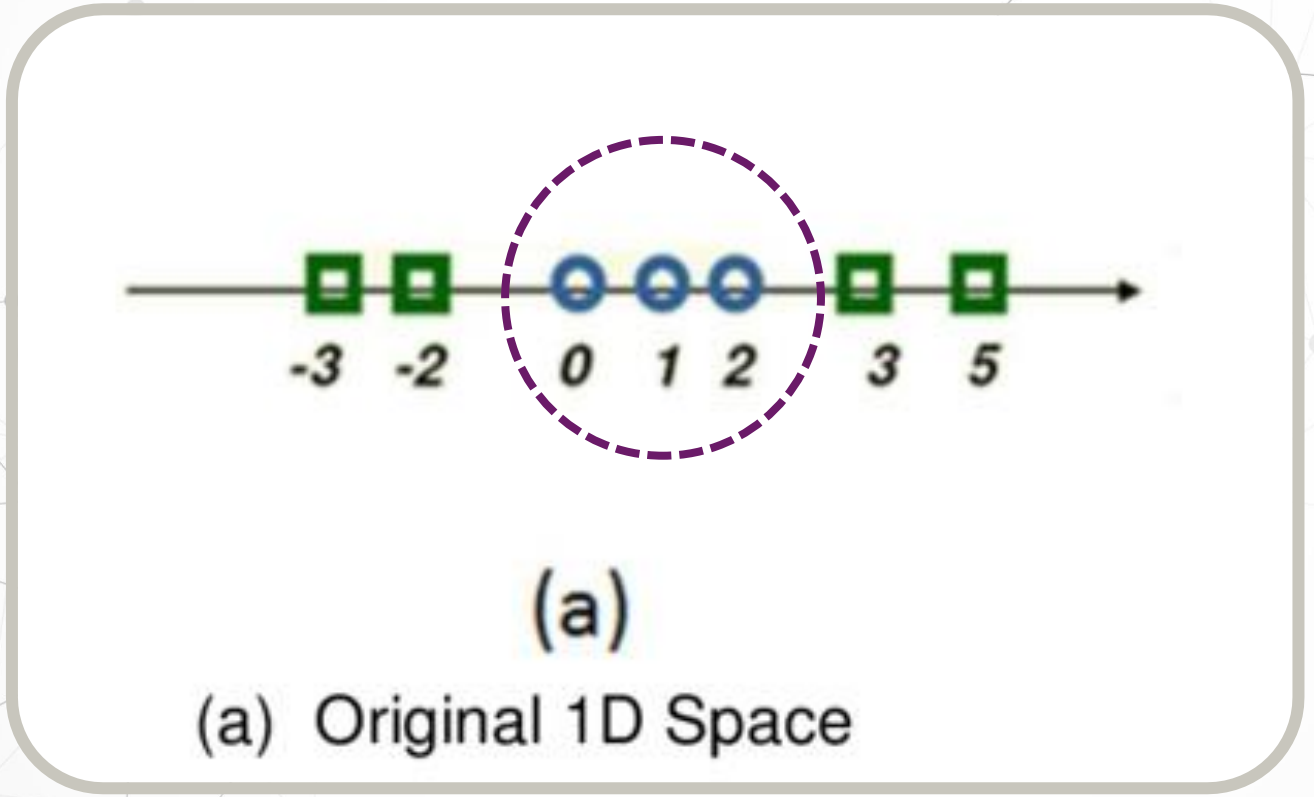
Support Vector Machines



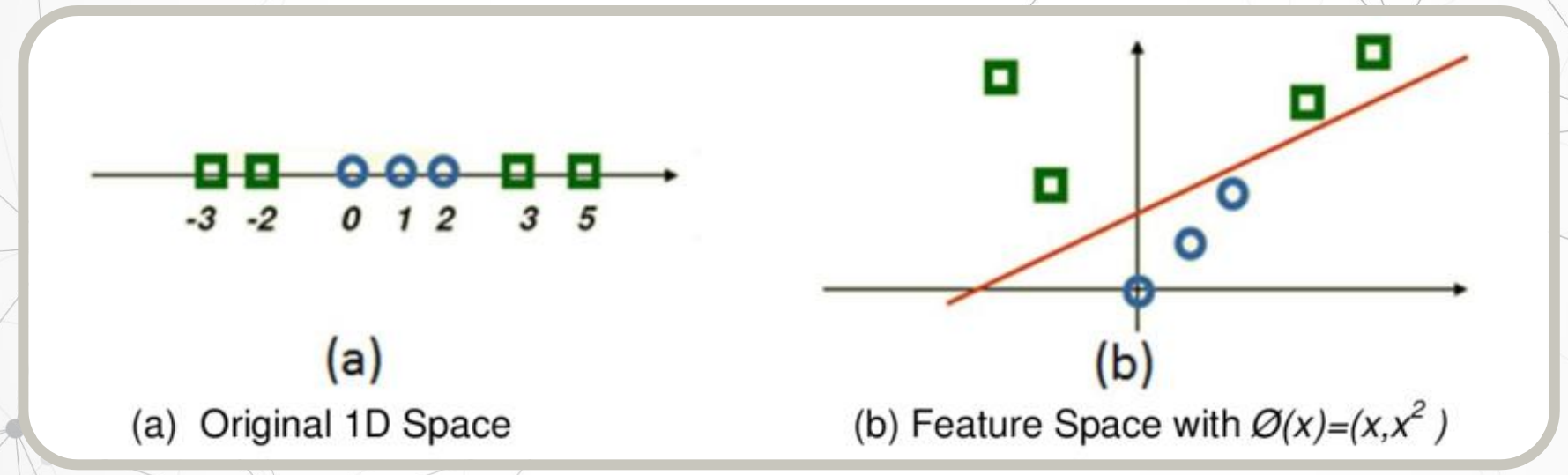
Support Vector Machines



Support Vector Machines



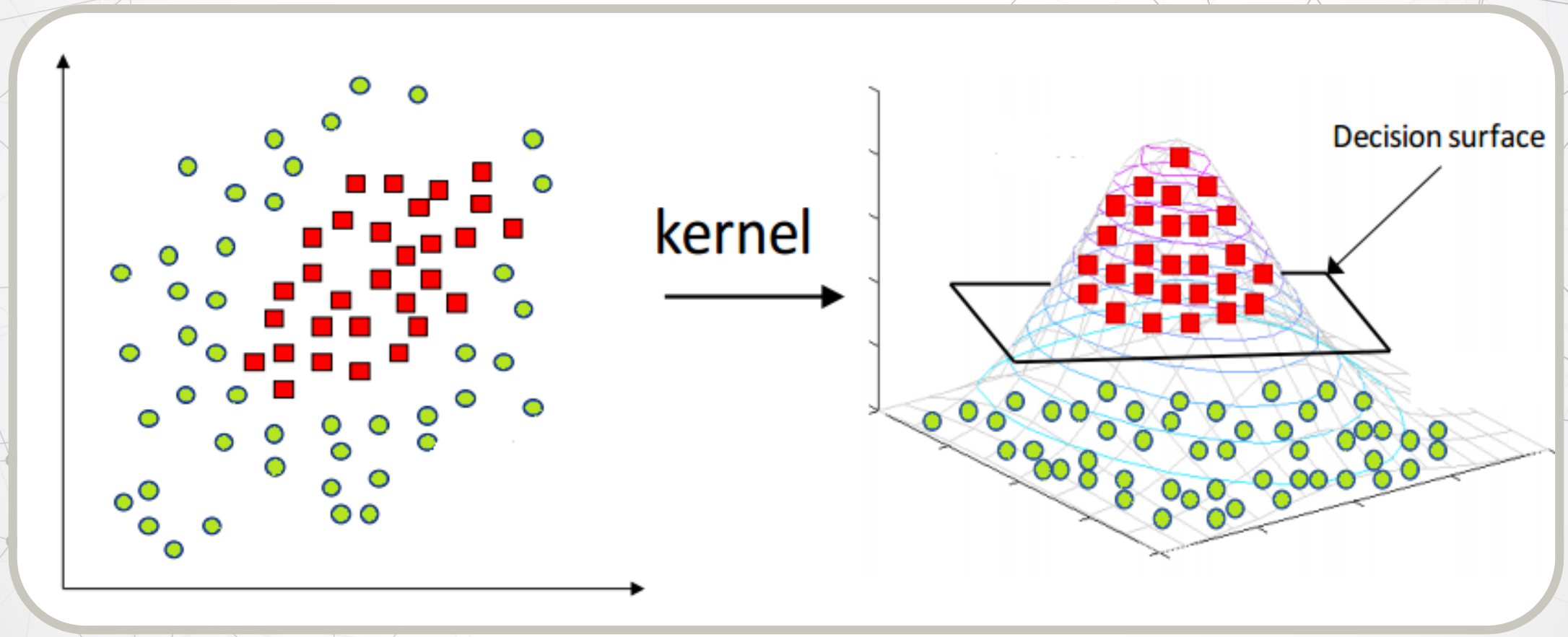
Support Vector Machines



(a) Original 1D Space

(b) Feature Space with $\emptyset(x)=(x,x^2)$

Support Vector Machines

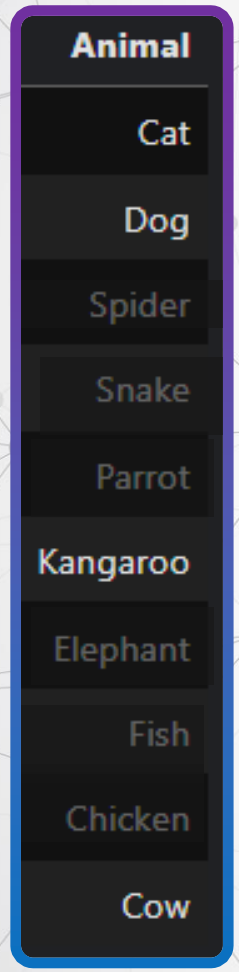


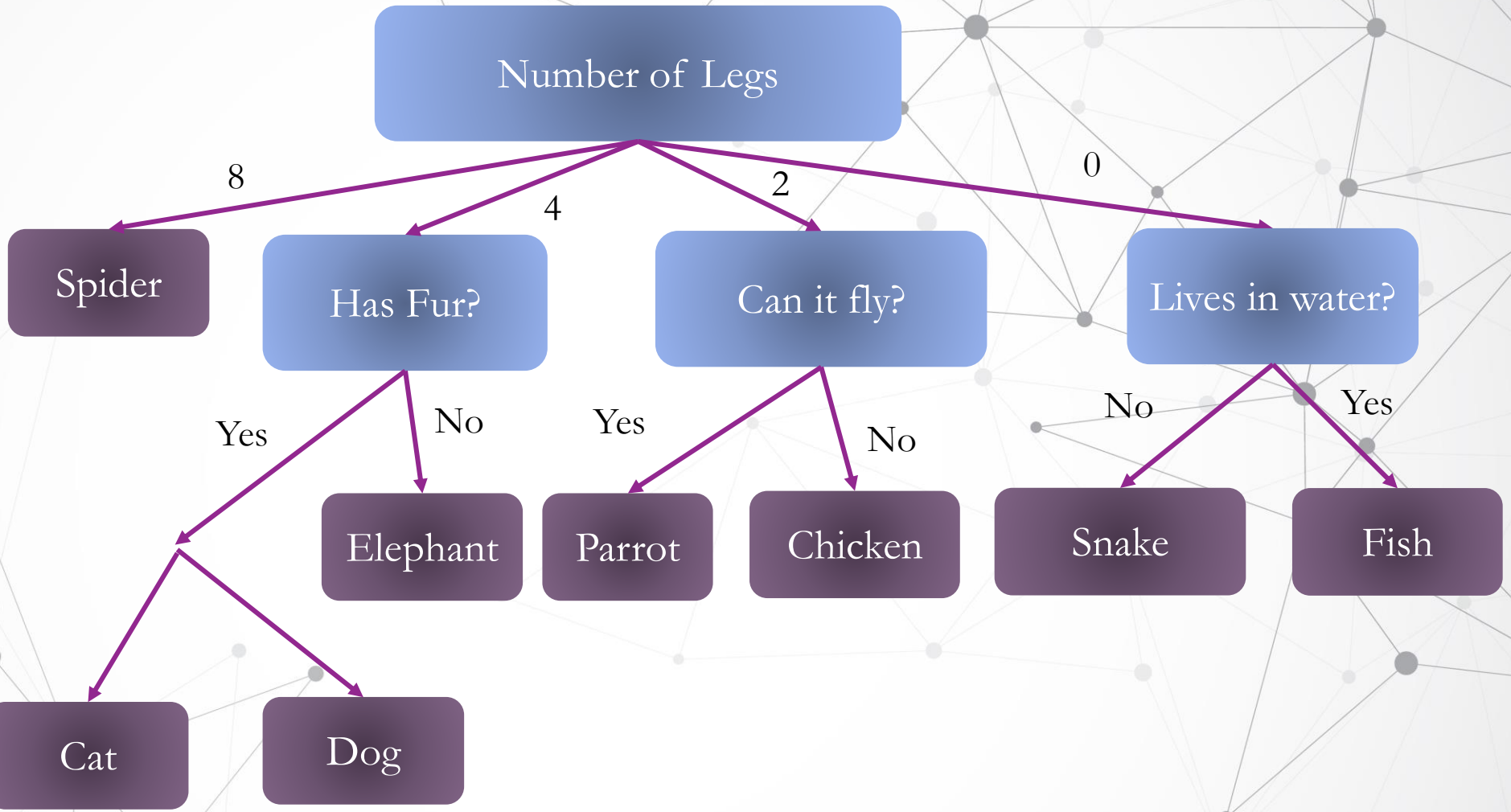
Decision Trees

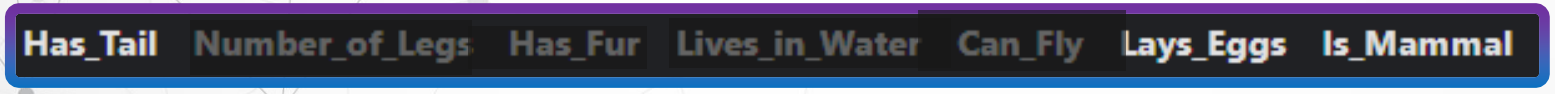
Has_Tail	Number_of_Legs	Has_Fur	Lives_in_Water	Can_Fly	Lays_Eggs	Is_Mammal	Animal
True	4	True	False	False	False	True	Cat
True	4	True	False	False	False	True	Dog
False	8	False	False	False	True	False	Spider
False	0	False	False	False	True	False	Snake
False	2	False	False	True	True	False	Parrot
True	2	True	False	False	False	True	Kangaroo
True	4	False	False	False	False	True	Elephant
False	0	False	True	False	True	False	Fish
False	2	False	False	False	True	False	Chicken
True	4	False	False	False	False	True	Cow

Decision Trees

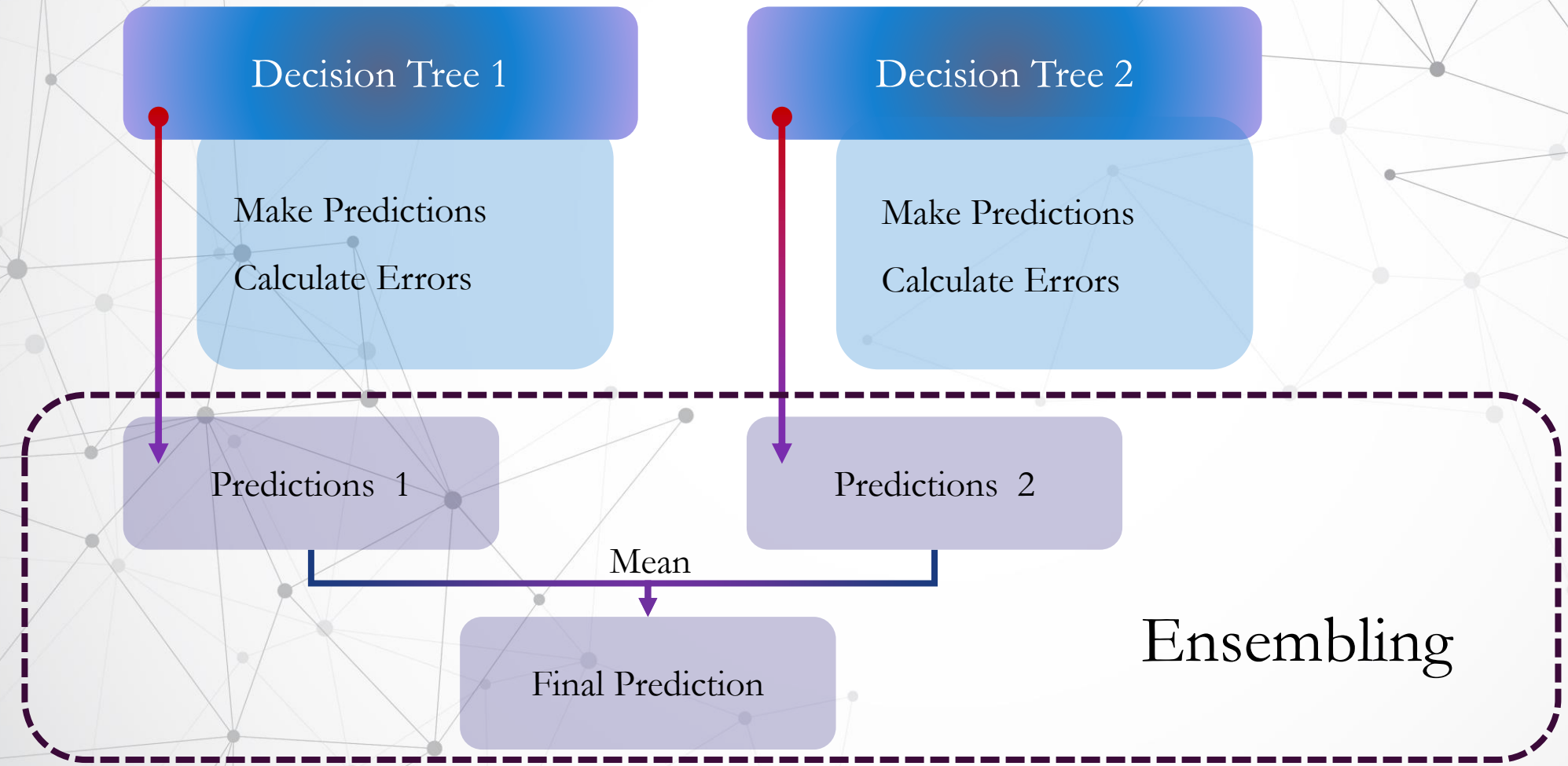
Has_Tail	Number_of_Legs	Has_Fur	Lives_in_Water	Can_Fly	Lays_Eggs	Is_Mammal	Animal
True	4	True	False	False	False	True	Cat
True	4	True	False	False	False	True	Dog
False	8	False	False	False	True	False	Spider
False	0	False	False	False	True	False	Snake
False	2	False	False	True	True	False	Parrot
True	2	True	False	False	False	True	Kangaroo
True	4	False	False	False	False	True	Elephant
False	0	False	True	False	True	False	Fish
False	2	False	False	False	True	False	Chicken
True	4	False	False	False	False	True	Cow


Animal
Cat
Dog
Spider
Snake
Parrot
Kangaroo
Elephant
Fish
Chicken
Cow

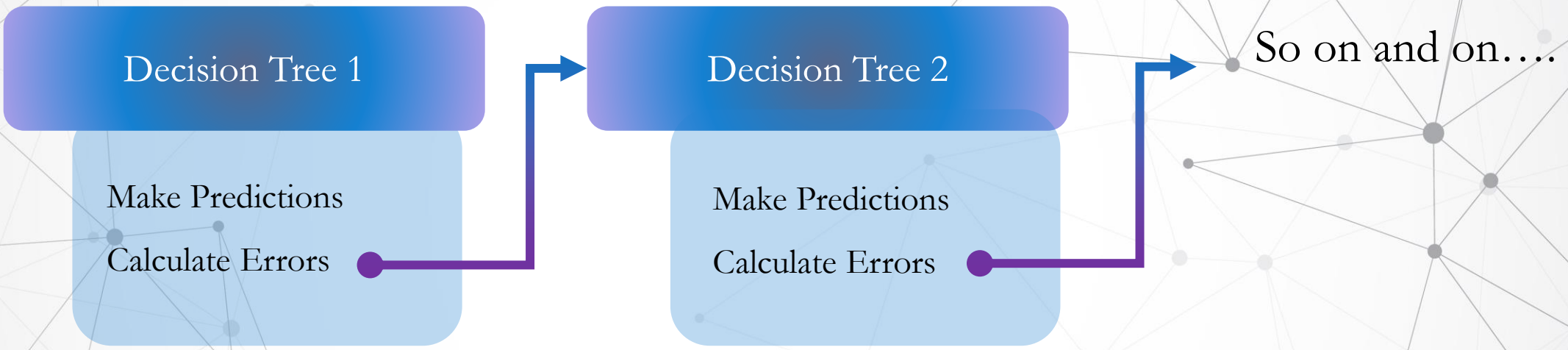

Number of Legs
8
4
2
0
Spider
Has Fur?
Can it fly?
Lives in water?
Yes
No
Yes
No
No
Yes
Elephant
Parrot
Chicken
Snake
Fish
Cat
Dog


Has_Tail
Number_of_Legs
Has_Fur
Lives_in_Water
Can_Fly
Lays_Eggs
Is_Mammal

Random Forest



XGBoost


Decision Tree 1
Make Predictions
Calculate Errors
Decision Tree 2
Make Predictions
Calculate Errors
So on and on….

Any Questions?