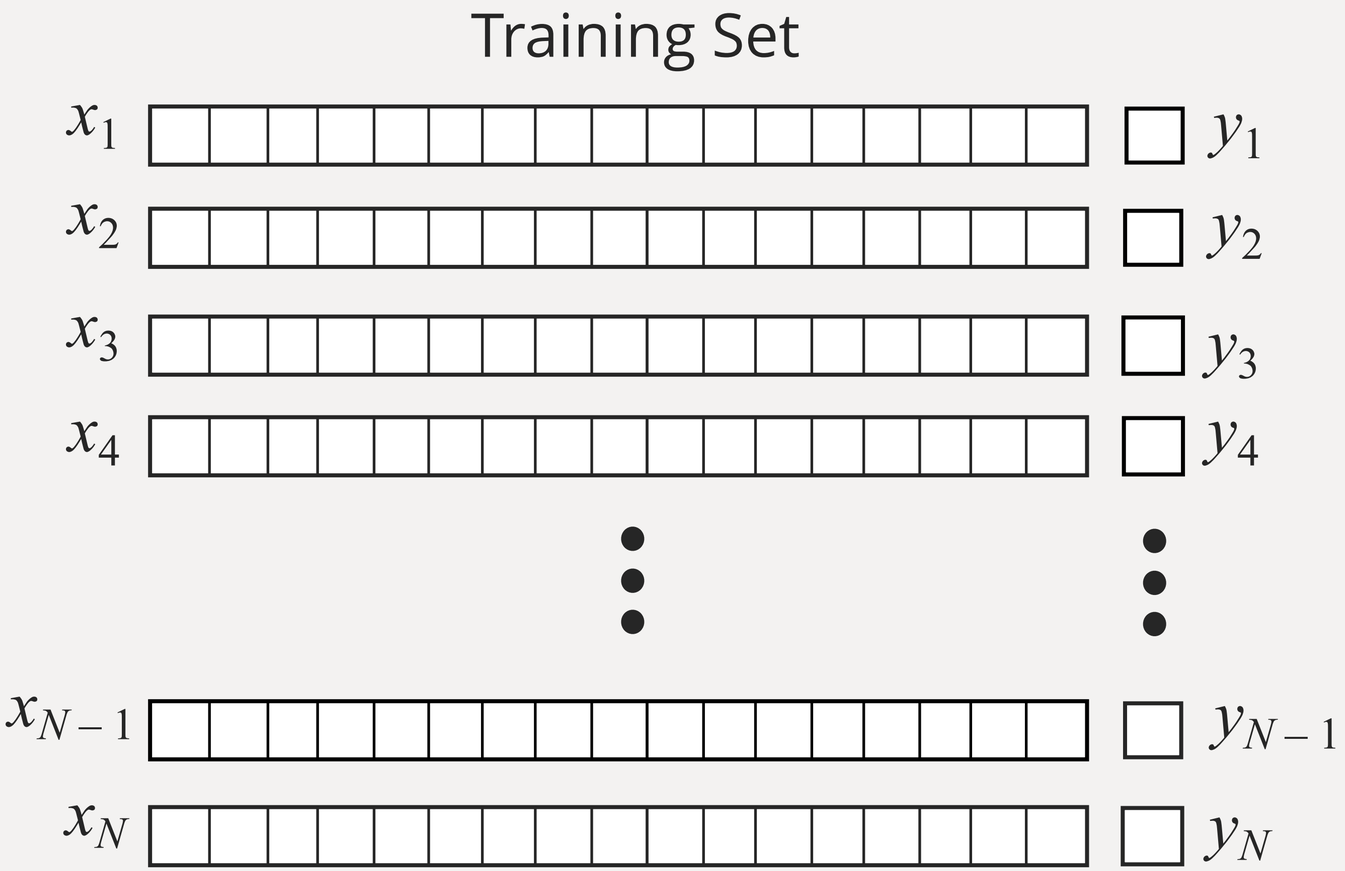
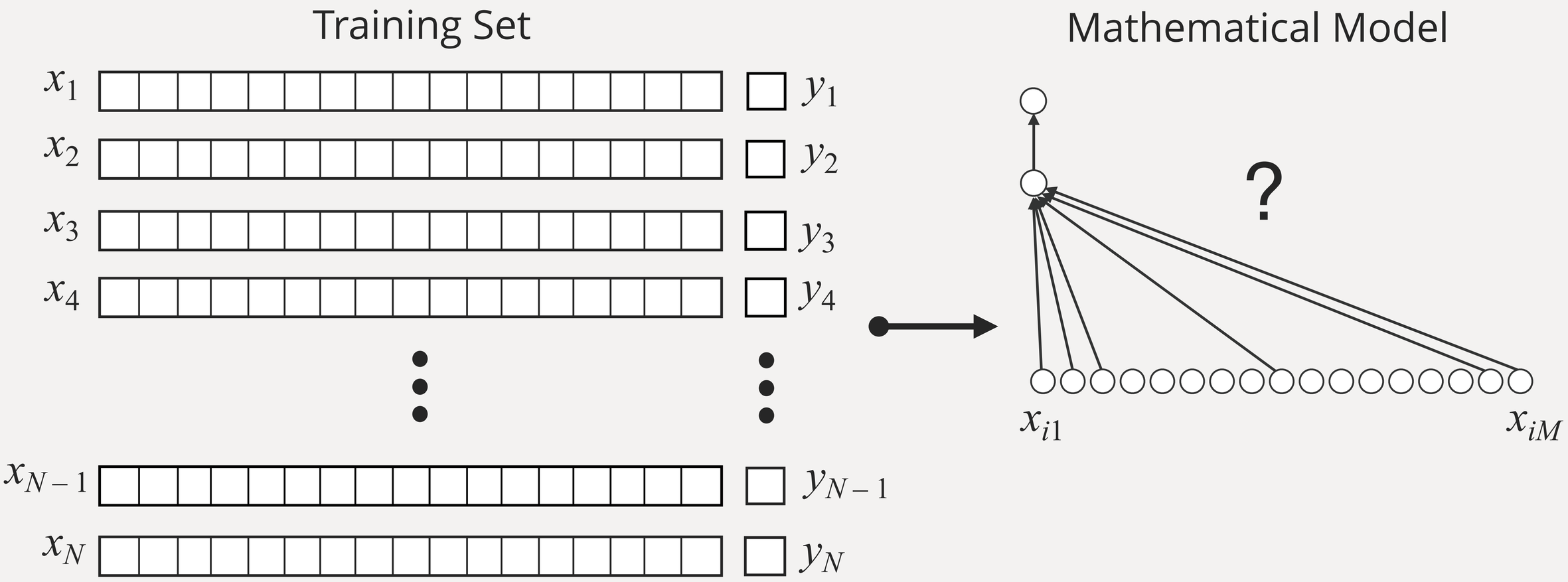
Logistic Regression

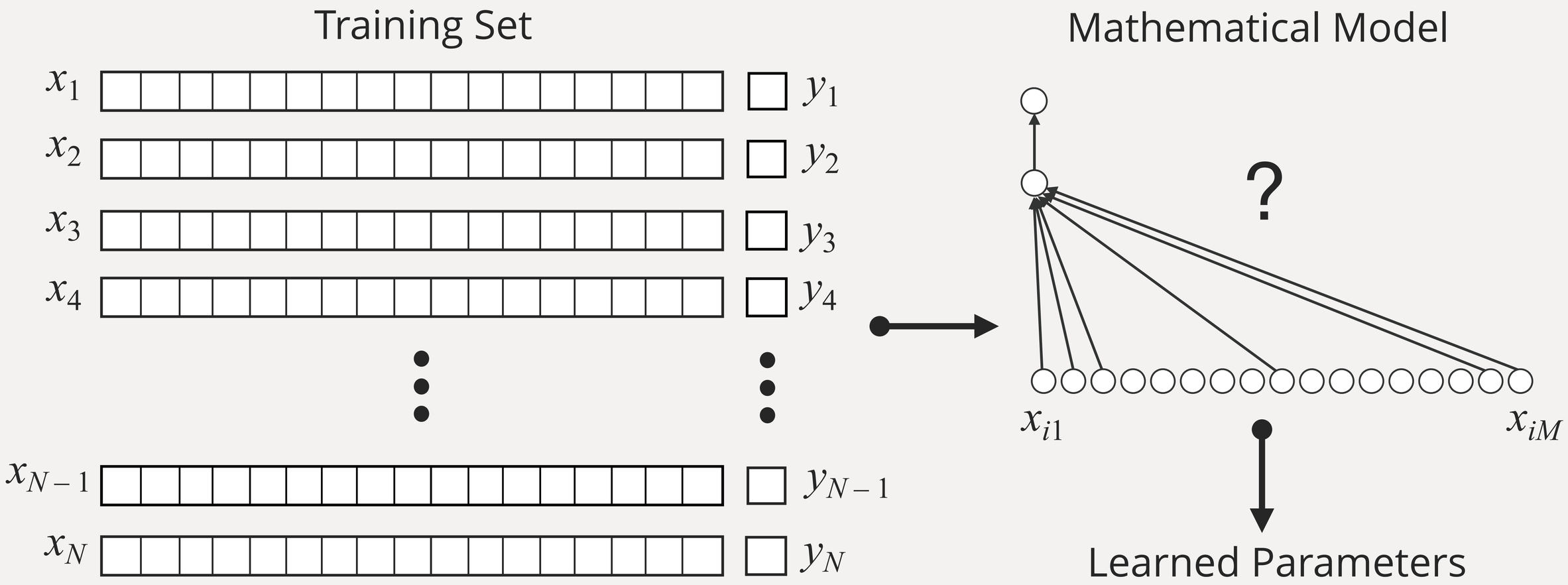
Learned Model Parameters



Learned Model Parameters



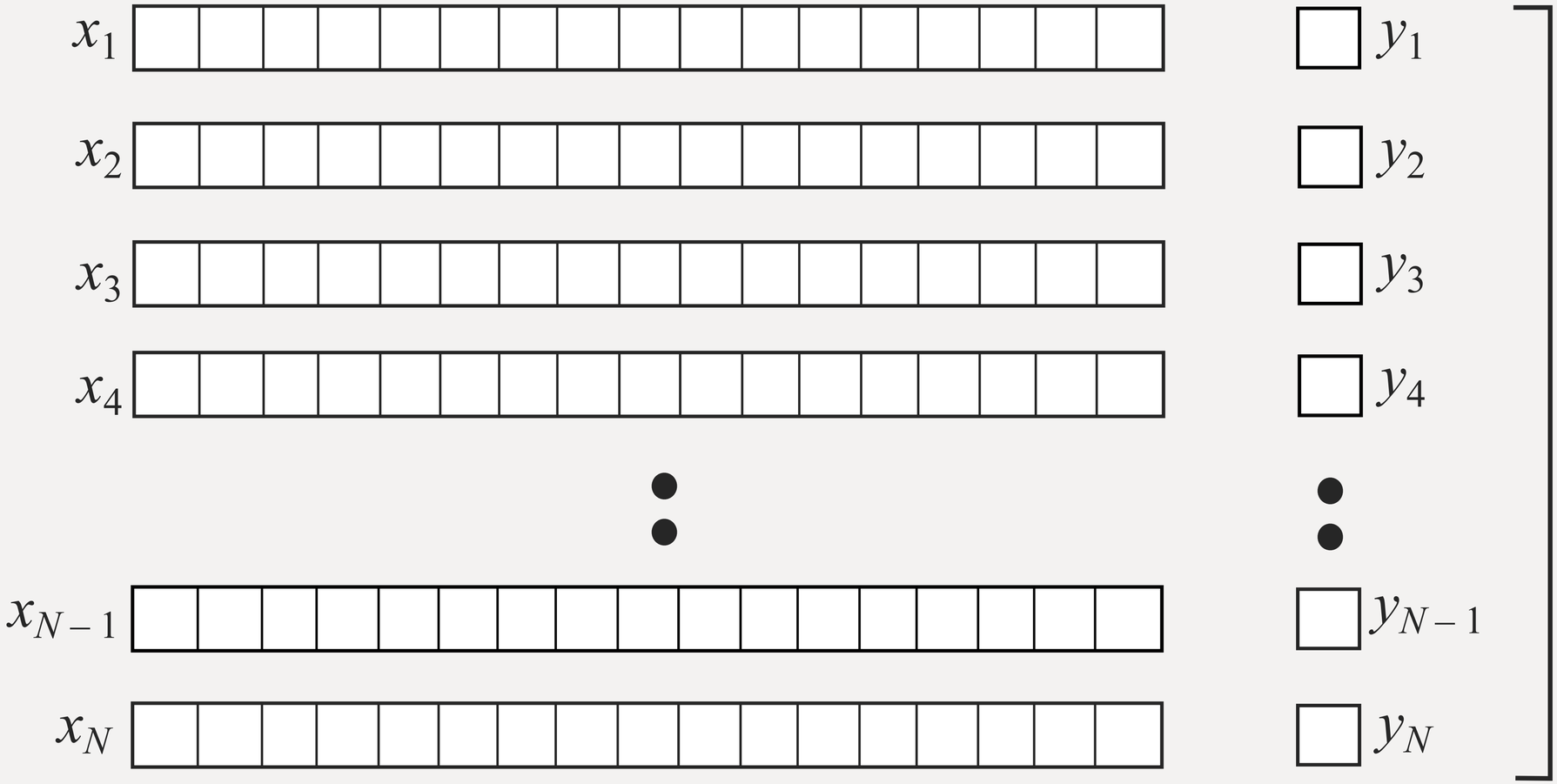
Learned Model Parameters



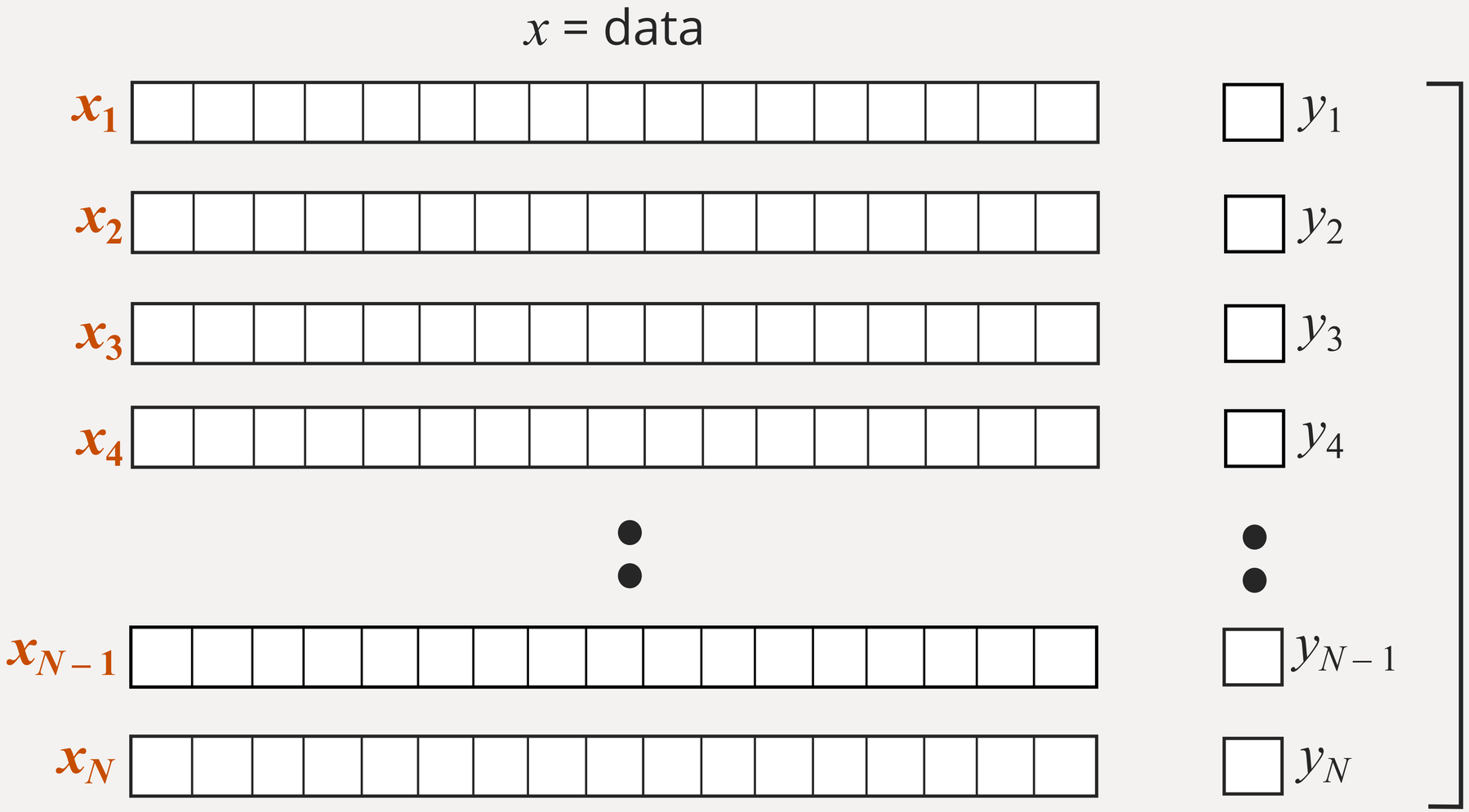
Learning infers what parameters of the model are consistent with training data

Logistic Regression

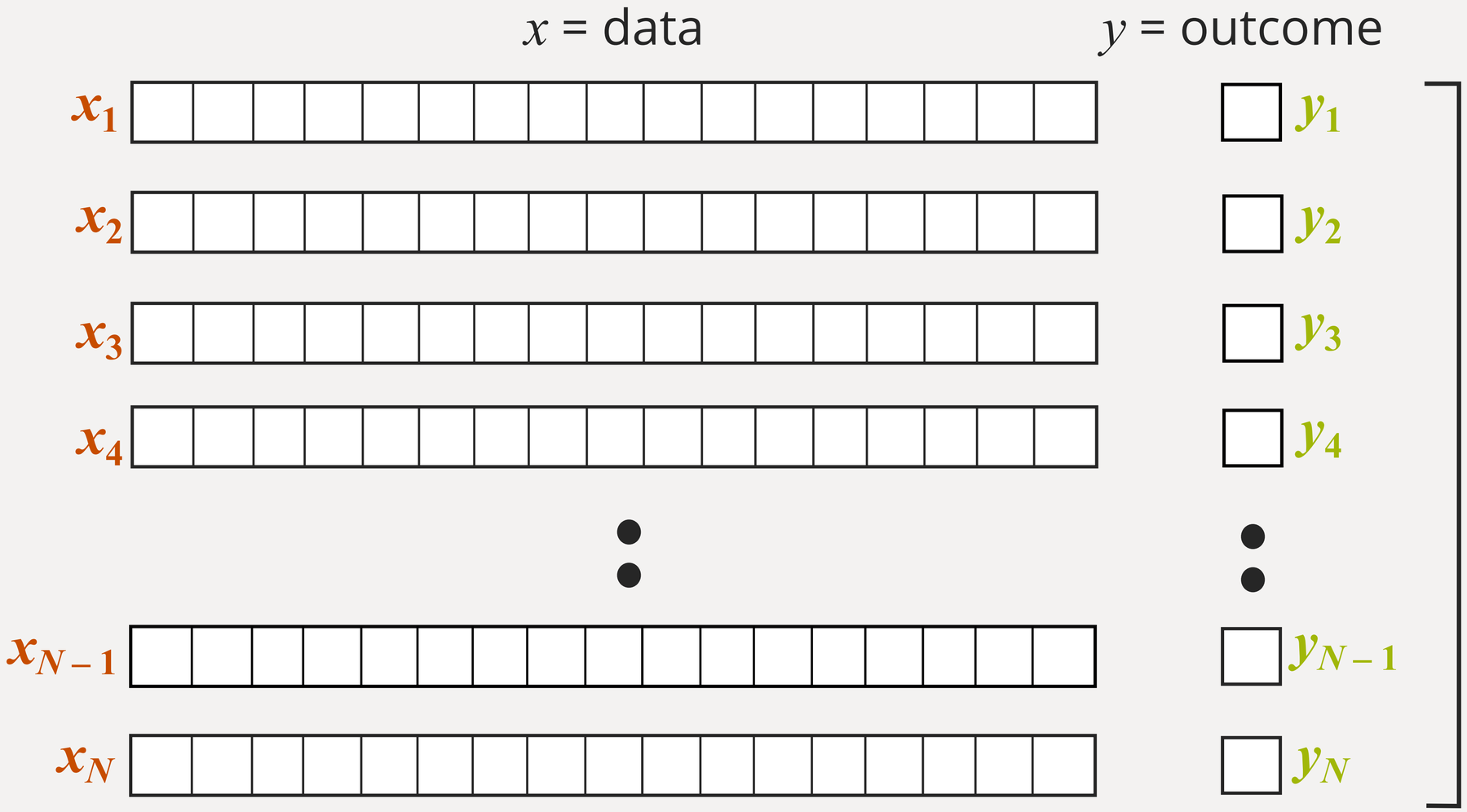
Training Set



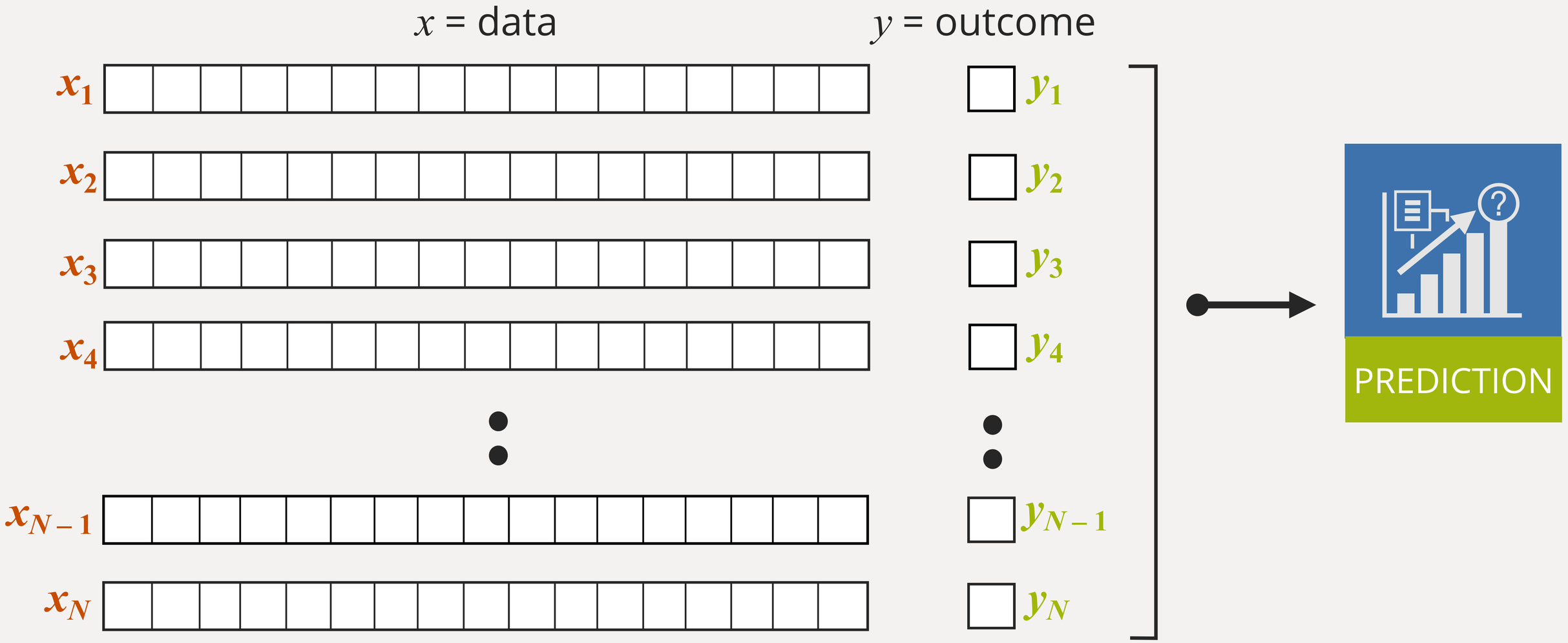
Training Set



Training Set



Training Set



Linear Predictive Model

Linear Predictive Model

$$(b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

bias

x_{i1} x_{i2} features of data x_{iM}

Linear Predictive Model

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

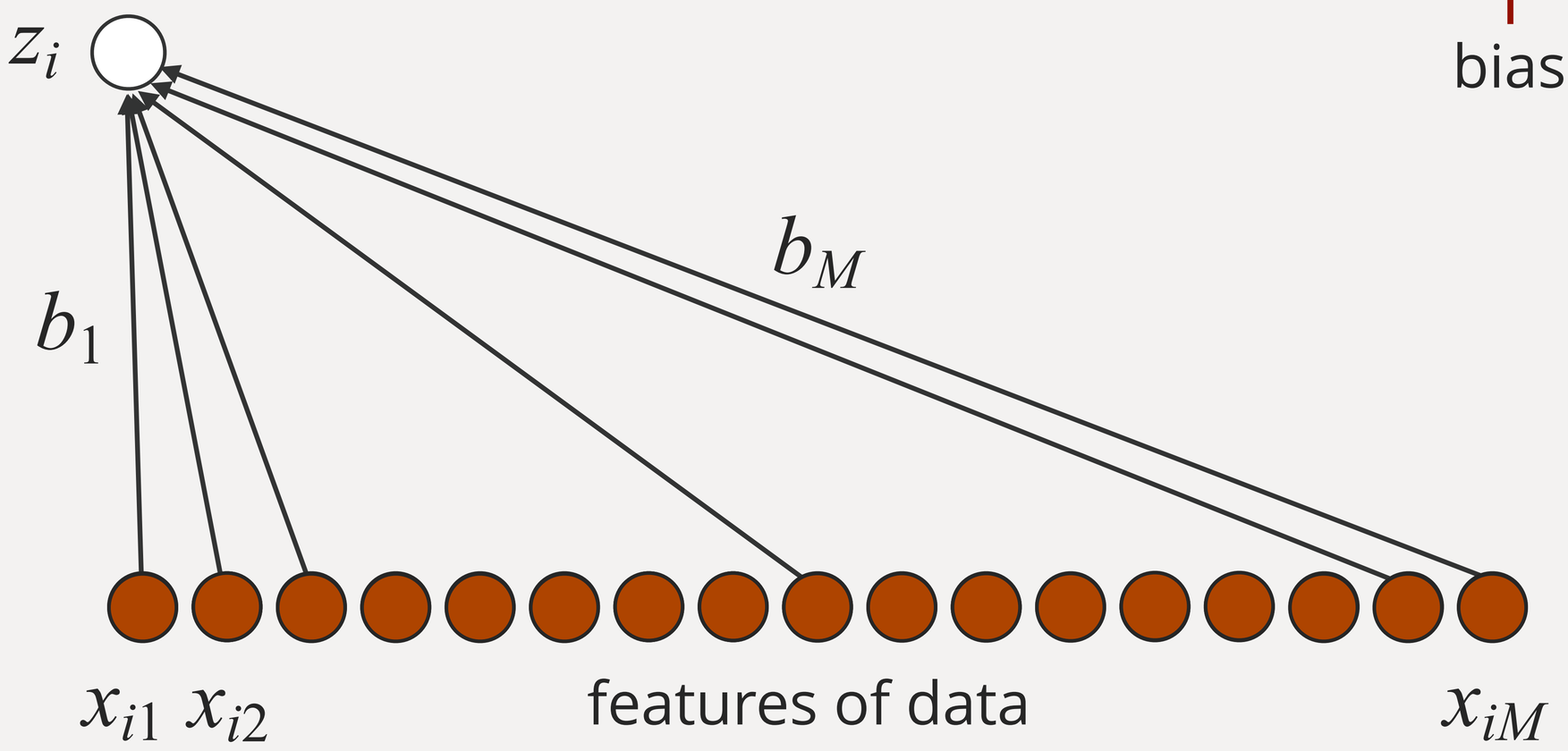
z_i

bias

x_{i1} x_{i2} features of data x_{iM}

Linear Predictive Model

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$



Will it Rain?

x_i = features for day i

features

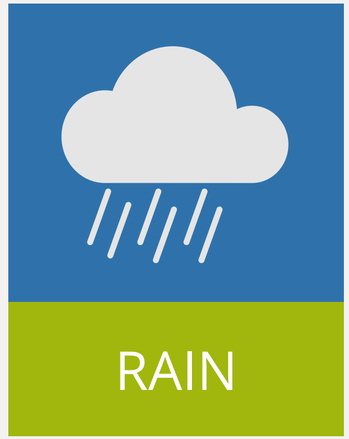
Cloud Cover	Humidity	Temperature	Air Pressure
0.5	80%	75	1.2
0.2	95%	83	1.3

outcome

Did it Rain
1
0

y_i = 1, yes

y_i = 0, no



$z_1 = (b_1 \text{ X } 0.5) + (b_2 \text{ X } 0.8) + (b_3 \text{ X } 75) + (b_4 \text{ X } 1.2) + b_0 \quad y_1 = 1$

$z_2 = (b_1 \text{ X } 0.2) + (b_2 \text{ X } 0.95) + (b_3 \text{ X } 83) + (b_4 \text{ X } 1.3) + b_0 \quad y_2 = 0$

$$p(y_i = 1 | x_i)$$

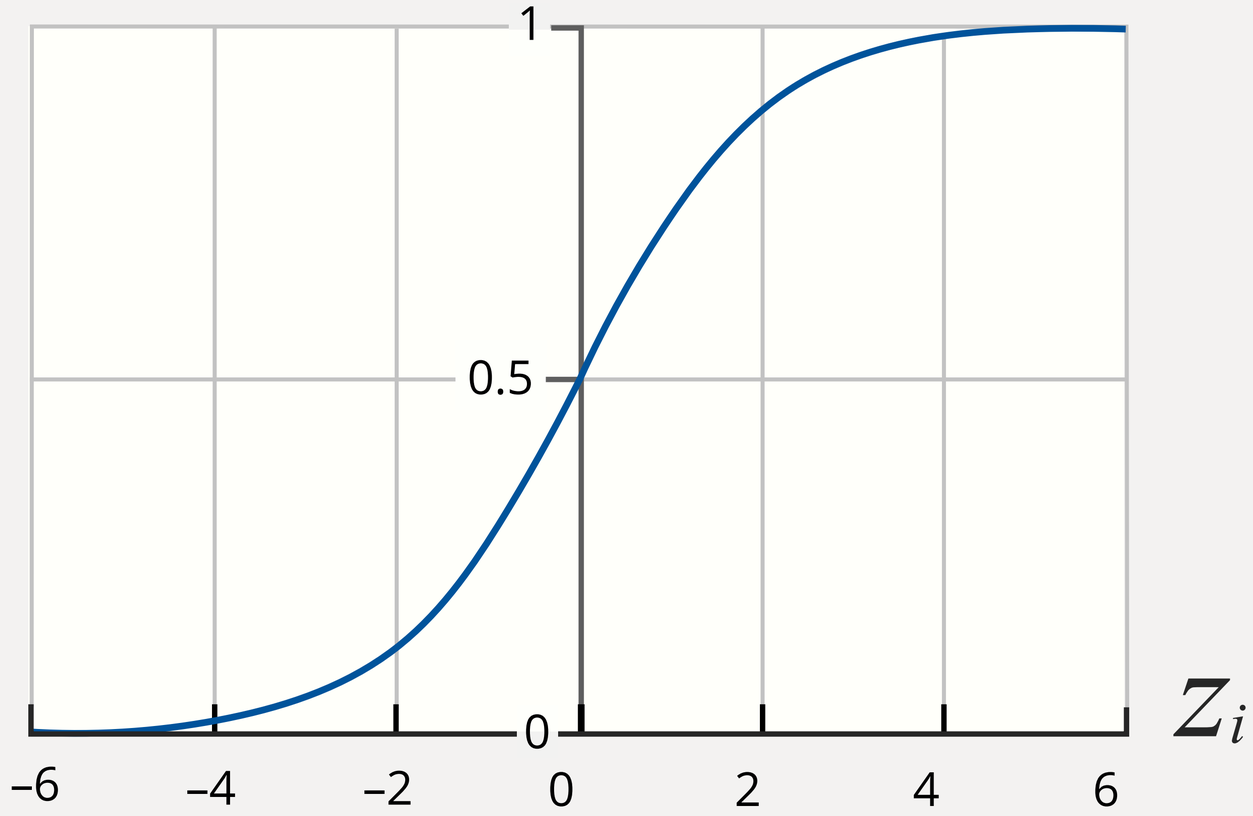
sigma

$$p(y_i = 1|x_i) = \sigma(z_i)$$

Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

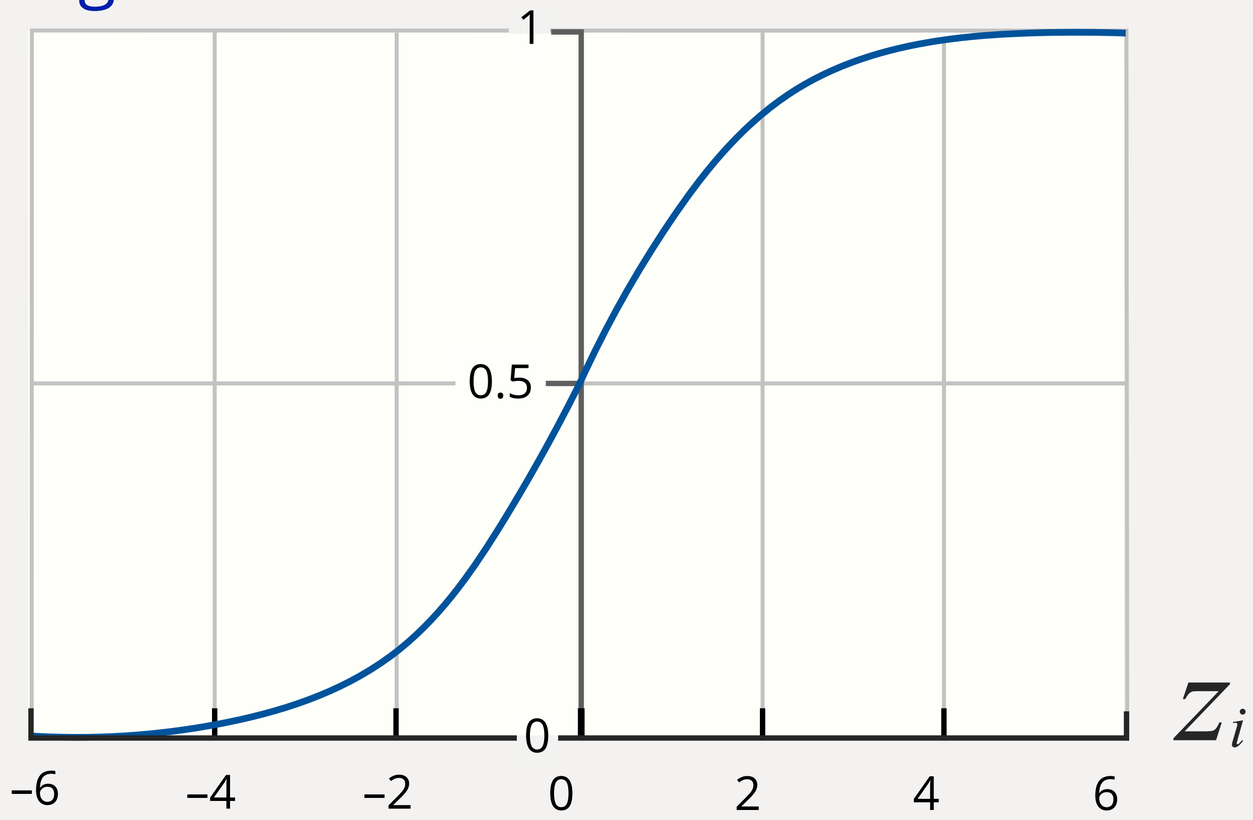
$$p(y_i = 1|x_i) = \sigma(z_i)$$



Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

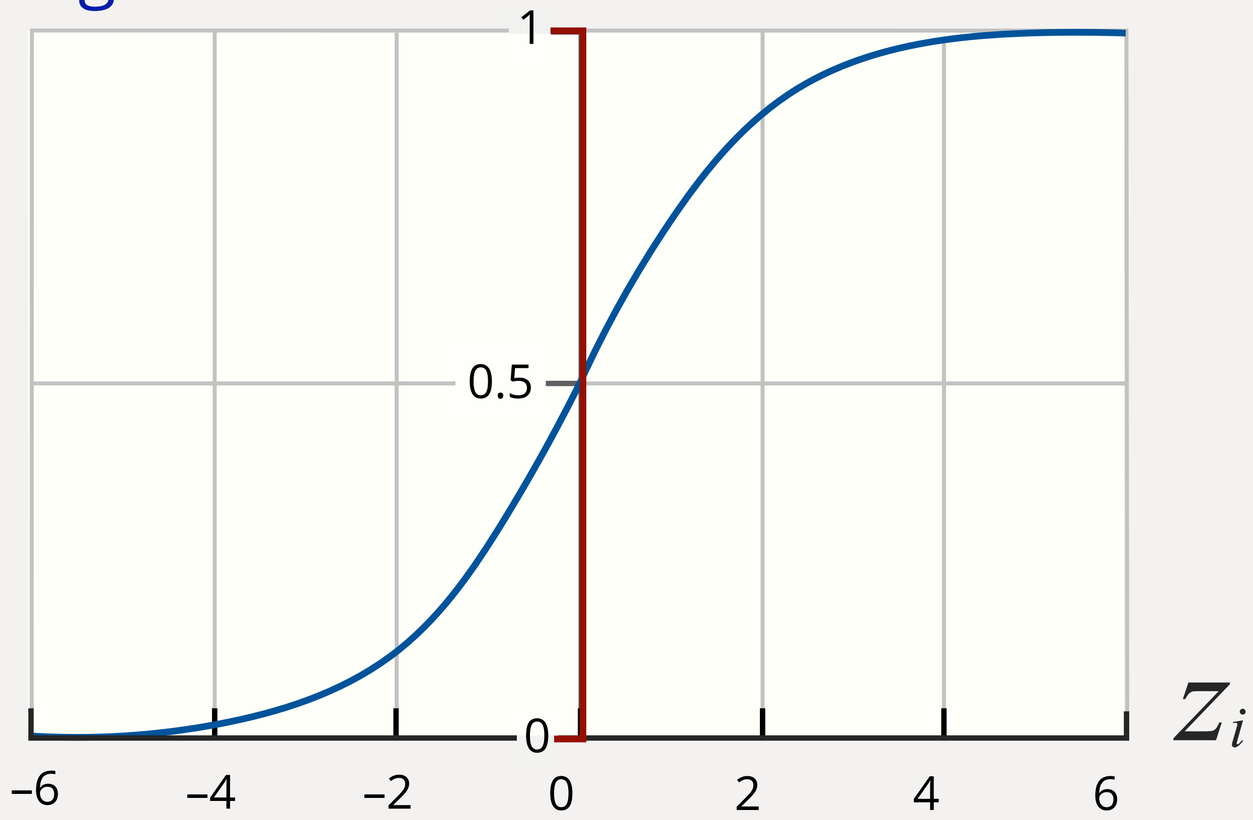
Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

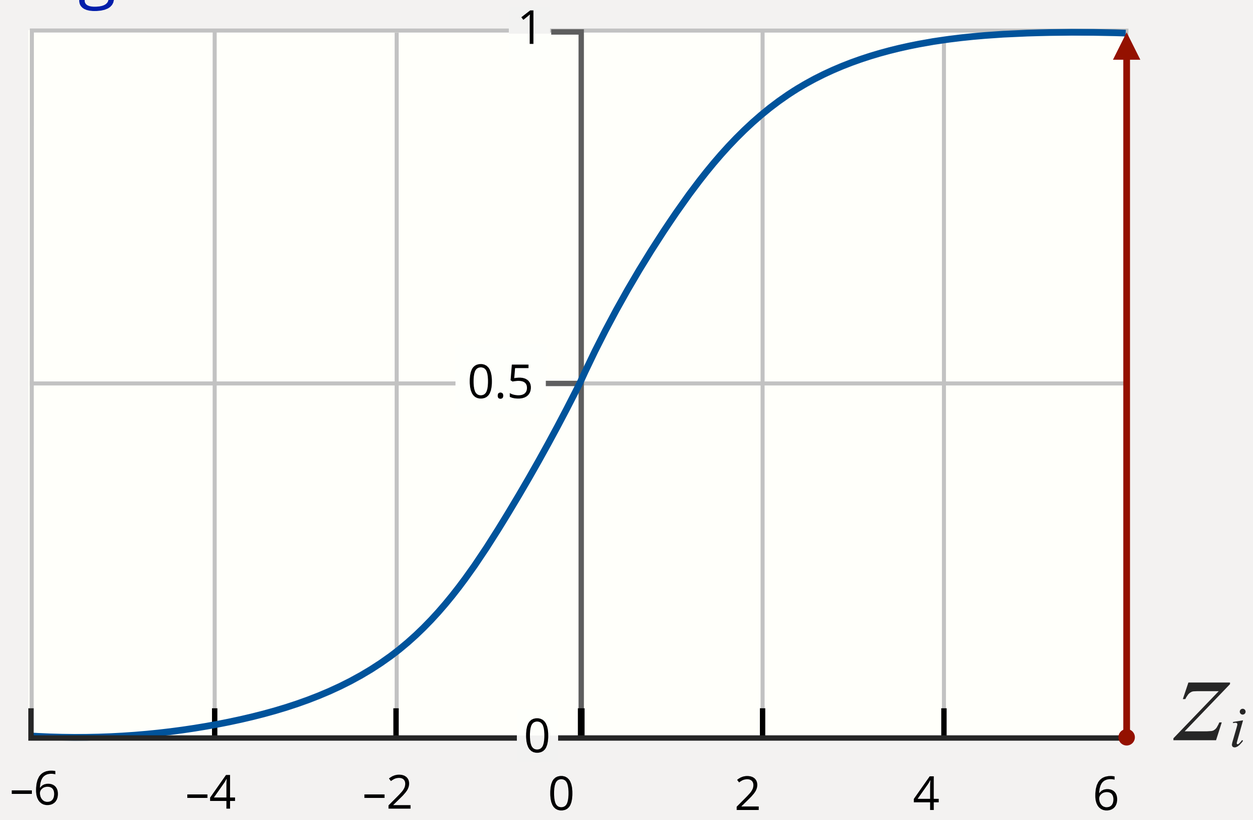
Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

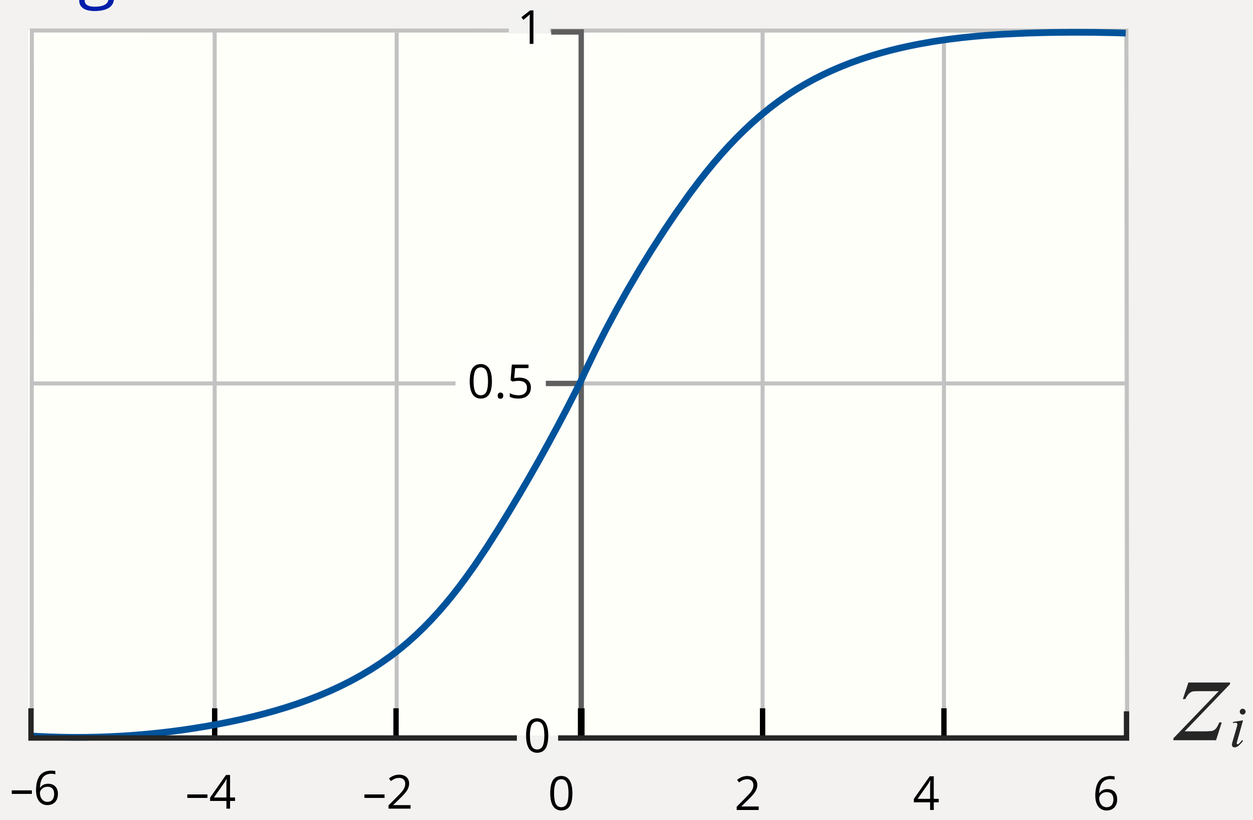
Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

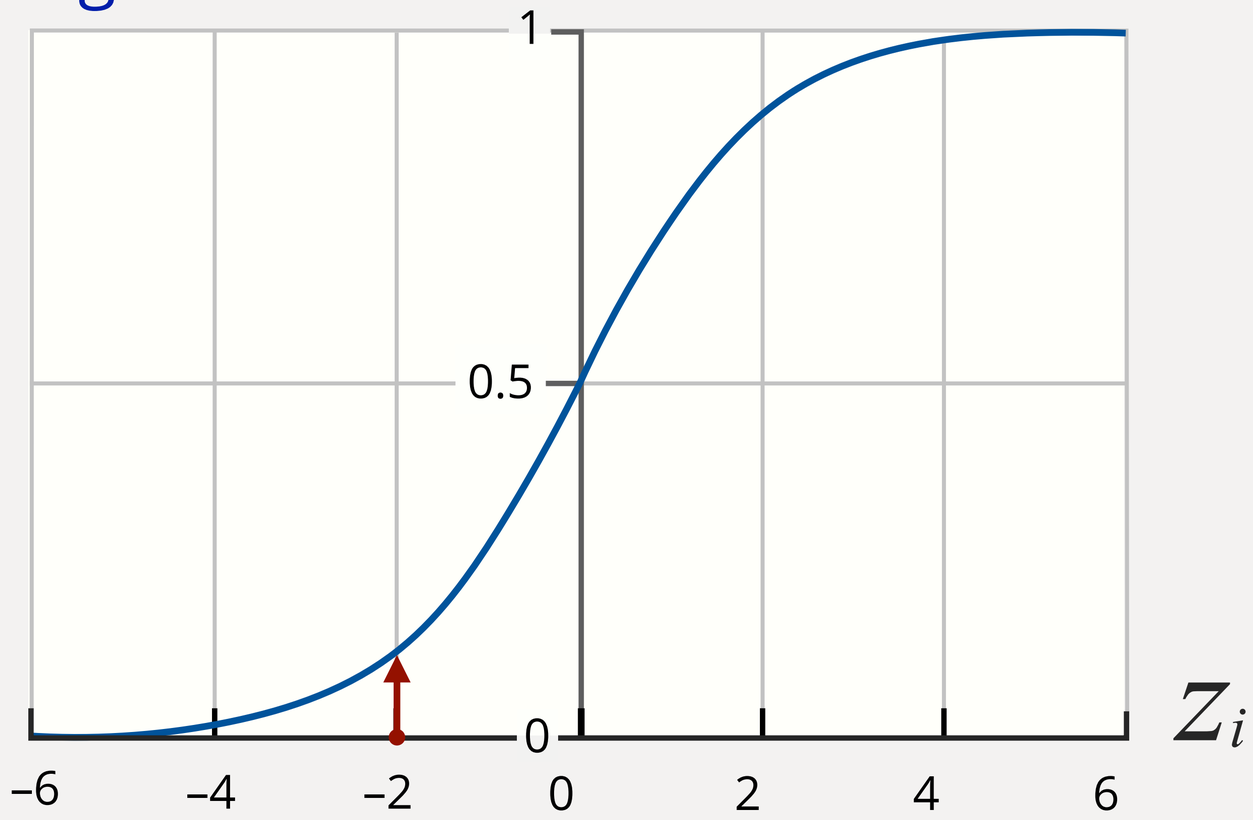
Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

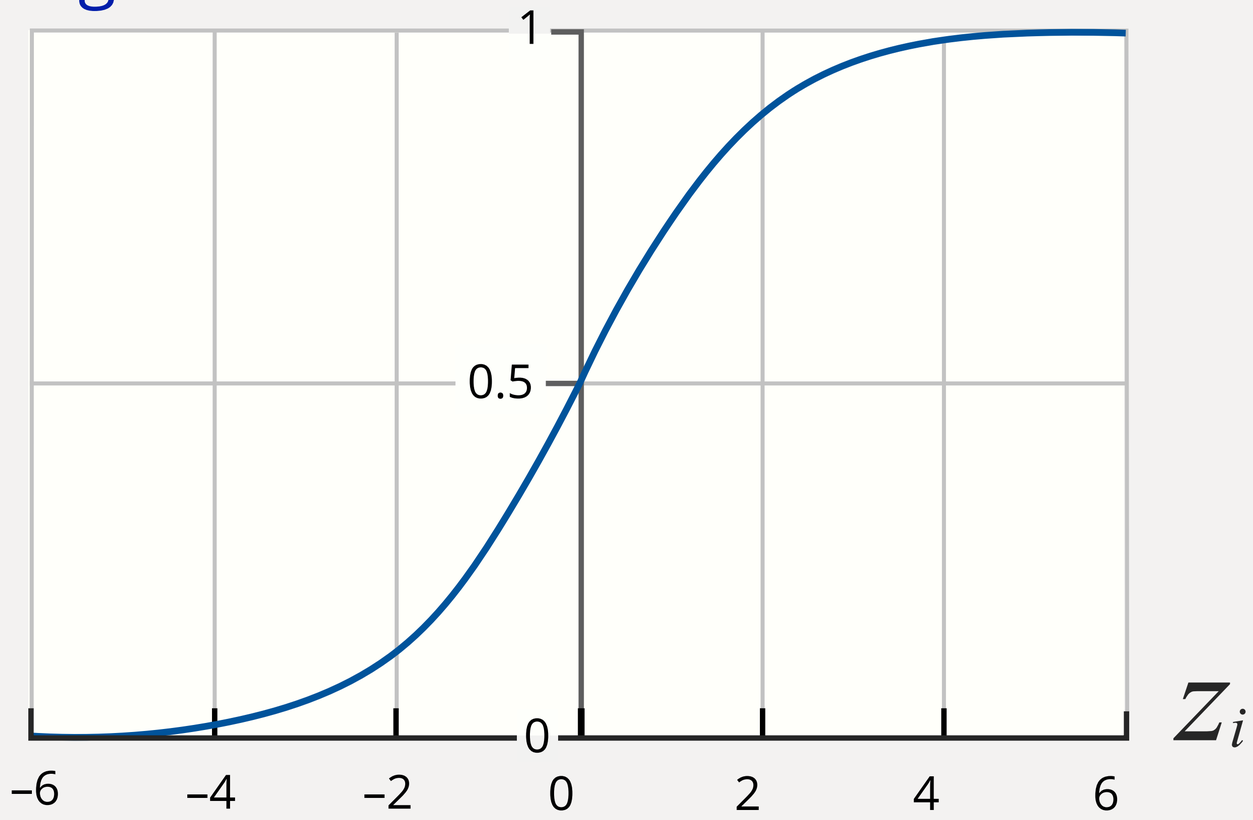
Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$

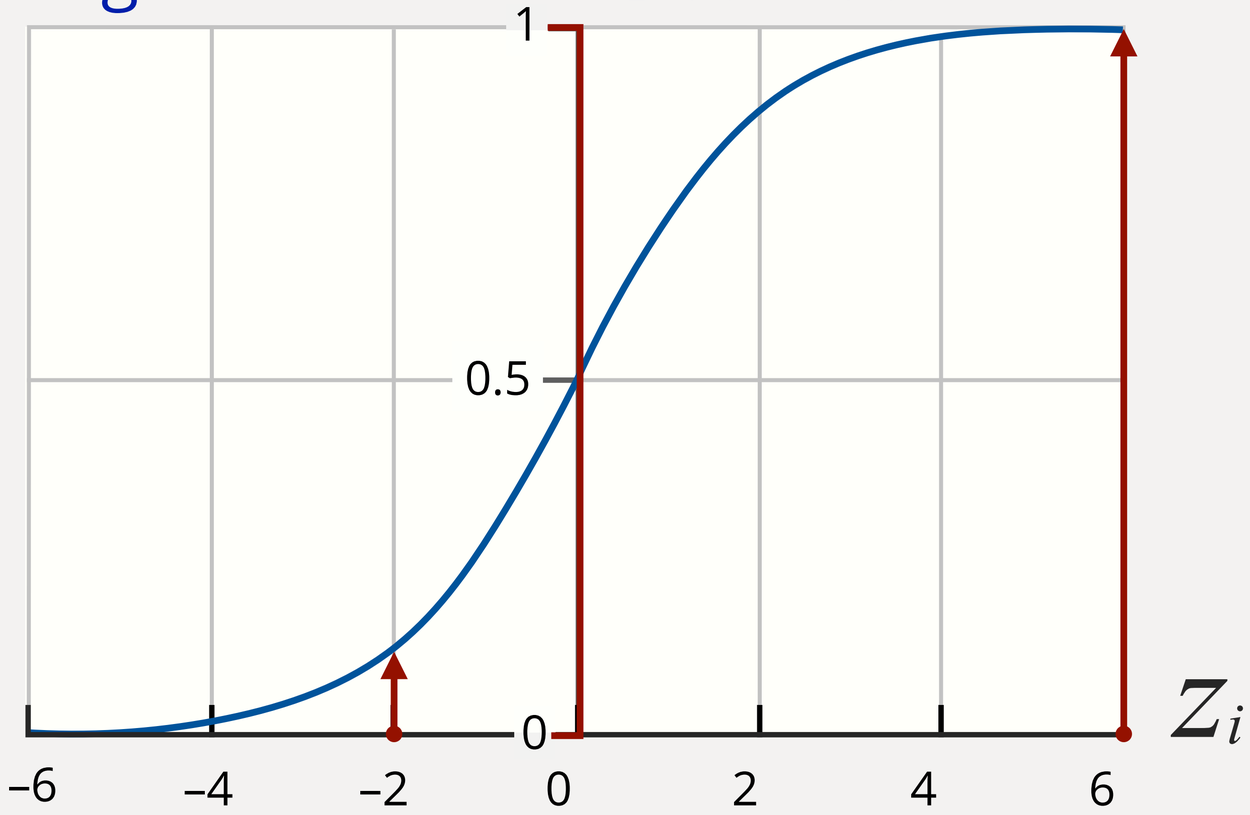


Sigmoid Function is a way to convert predictions to a probabilistic perspective

Convert to a Probability

$$z_i = (b_1 \times x_{i1}) + (b_2 \times x_{i2}) + \cdots + (b_M \times x_{iM}) + b_0$$

Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



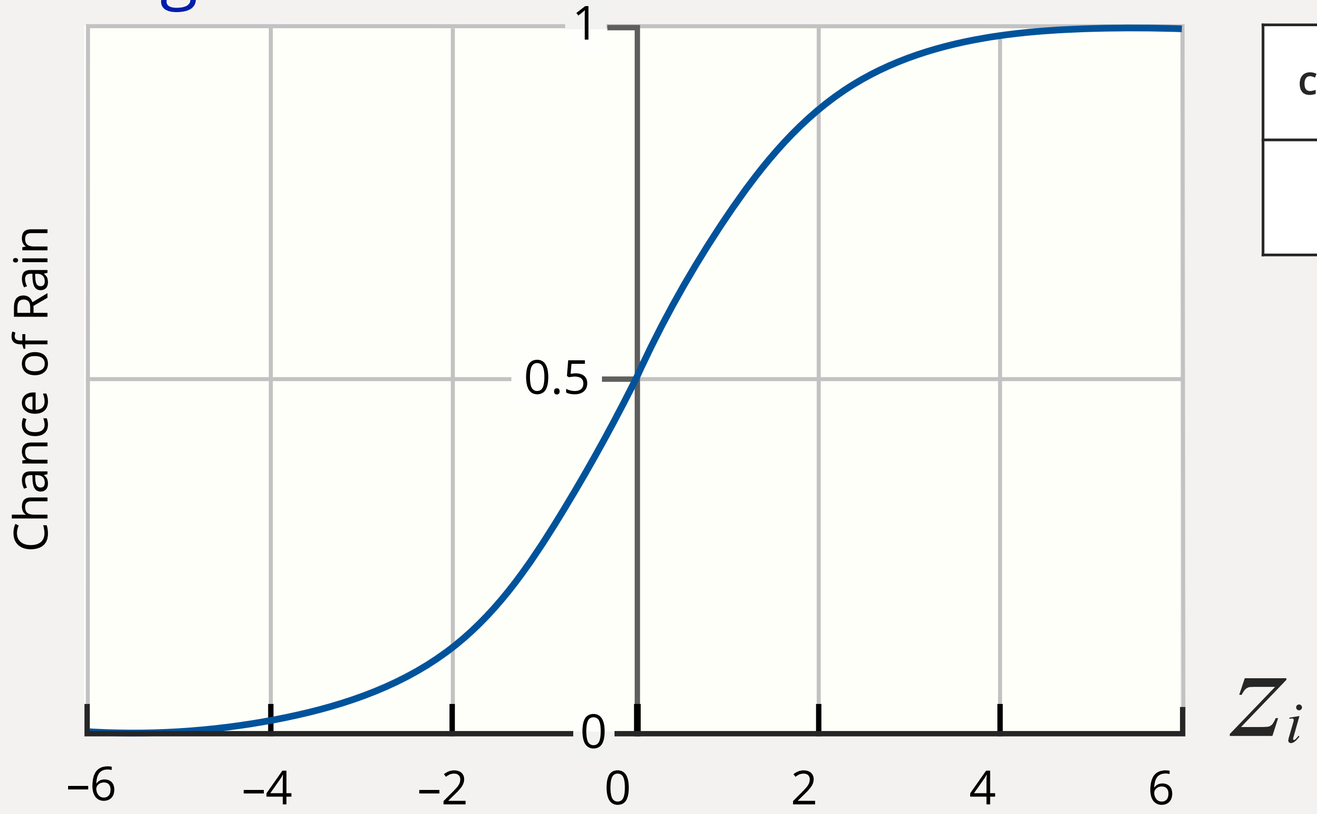
Outcome of Z

- Z_i = Large and positive indicates y_i = 1 is likely
- Z_i = Large and negative indicates y_i = 0 is likely

Convert to a Probability

$$z_i = (b_1 \times 0.5) + (b_2 \times 0.8) + (b_3 \times 75) + (b_4 \times 1.2) + b_0$$

Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



features

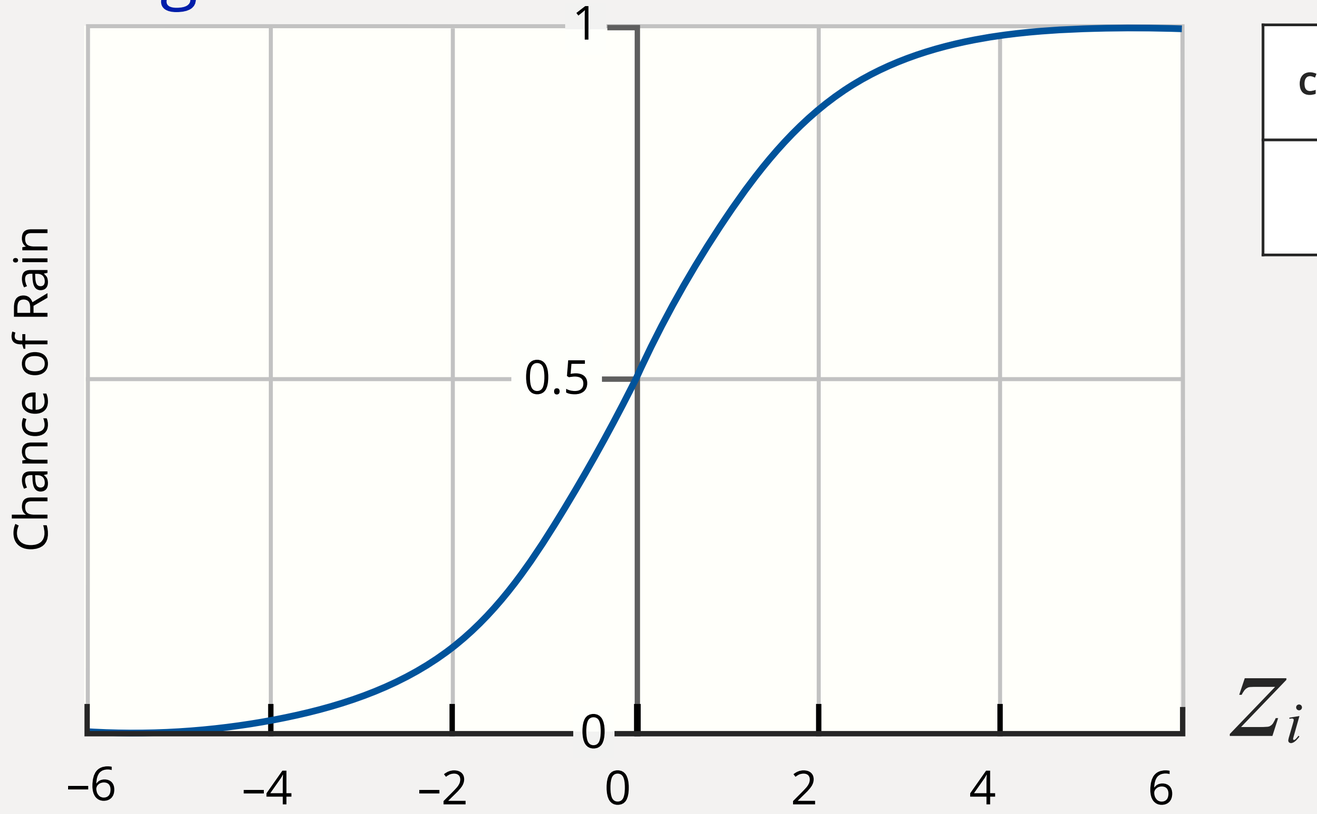
Cloud Cover	Humidity	Temperature	Air Pressure
0.5	80%	75	1.2

b parameters tell us how important data variables are to the prediction

Convert to a Probability

$$z_i = (b_1 \times 0.5) + (b_2 \times 0.8) + (b_3 \times 75) + (b_4 \times 1.2) + b_0$$

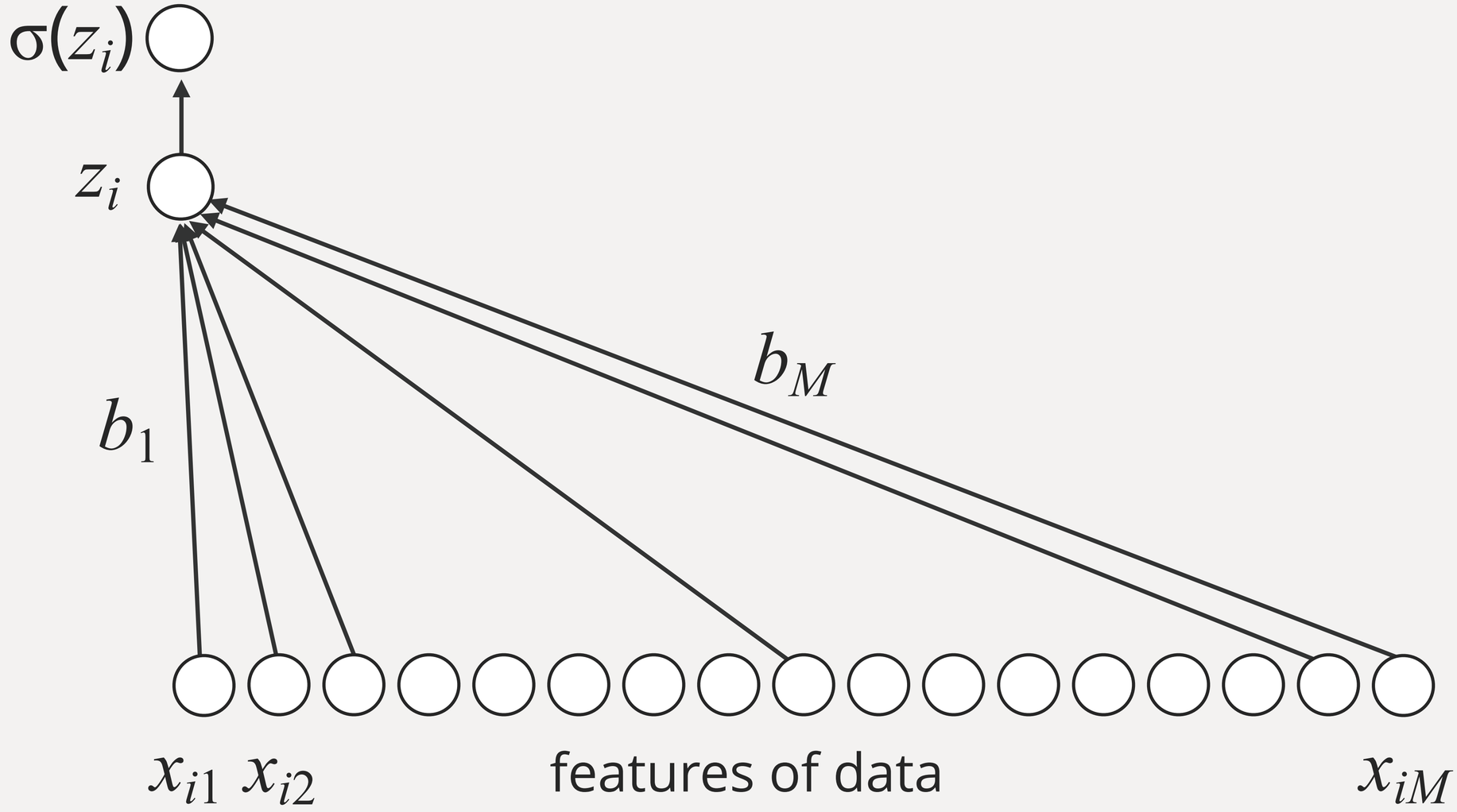
Sigmoid Function $p(y_i = 1|x_i) = \sigma(z_i)$



features

Cloud Cover	Humidity	Temperature	Air Pressure
0.5	80%	75	1.2

Logistic Regression



Learned Model Parameters