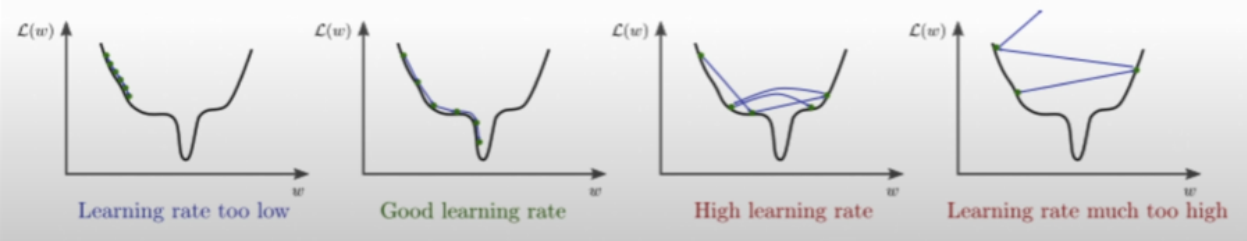
Linear Regression

gradient descent: optimization algorithm that is used to find the minimum of a function by adjusting its parameters



Steps

Training:

- Initialize weight as zero
- Initialize bias as zero

Given a data point:

- Predict result by using y = wx + b
- calculate error
- use gradient descent to figure out new weight and bias values
- repeat n times

$$\hat{y} = wx + b \longleftrightarrow y_{pred} = wX + b$$

$$X = [x_1 \quad x_2 \quad \dots \quad x_n]$$

$$wX = [wx_1 \quad wx_2 \quad \dots \quad wx_n]$$

$$y_{pred} = [wx_1 + b \quad wx_2 + b \quad \dots \quad wx_n + b]$$