

# STOCHASTIC GRADIENT DESCENT

1. Shuffle Observations.
2. For each observation:

$$W' = W - \eta \Delta \text{Loss}(W, x_i, y_i)$$

Diagram illustrating the stochastic gradient descent update formula:

- $W'$ : updated parameter (indicated by an upward arrow)
- $W$ : parameter (indicated by an upward arrow)
- $\eta$ : Learning Rate (indicated by an upward arrow)
- $\Delta$ : Gradient (indicated by an upward arrow)
- $\text{Loss}(W, x_i, y_i)$ : Loss function, where  $W$  is the Parameter (indicated by an upward arrow)
- $x_i$  and  $y_i$ : Individual  $x$  and  $y$  value (indicated by downward arrows)

3. Repeat until minimum value achieved.