

# **1. Stochastic gradient descent**

$\eta$  - learning rate  $w$  - weights vector  $x_i$  - one-shot input data vector