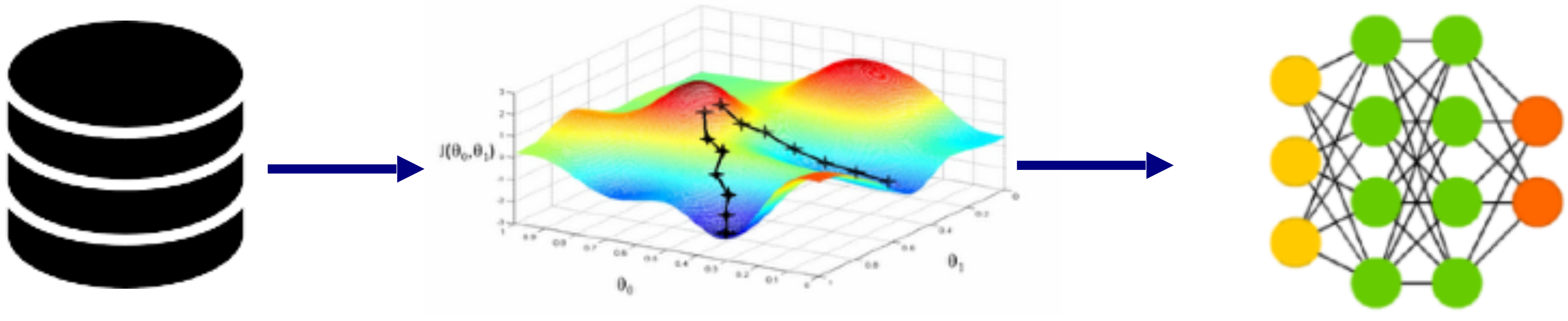
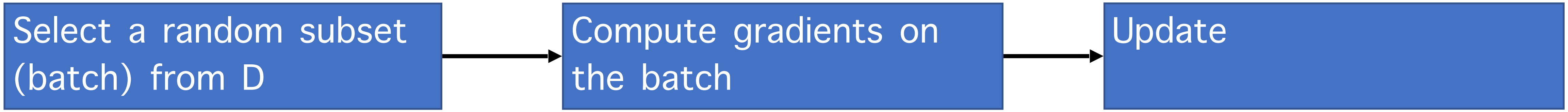
# Stochastic Gradient Descent (SGD)



Repeat ..



Select a random subset (batch) from D

Compute gradients on the batch

Update

.. until model convergence (i.e., many epochs)

# Differential Privacy (Weaker Notion)

$X$: The data *universe*.

$D \subset X$: The dataset (one element per person)

***Definition***: An algorithm $M$ is $(\boldsymbol{\epsilon}, \boldsymbol{\delta})$**-differentially private** if for all pairs of neighboring datasets $D, D'$, and for all outputs x:

$$\Pr[M(D) = x] \leq \exp(\boldsymbol{\epsilon}) * \Pr[M(D') = x] + \boldsymbol{\delta}$$

**Quantifies information leakage** **Allows for a small probability of failure**