

Ex: 4:

on-line learning (or) incremental learning

Batch GD \rightarrow See all data \rightarrow update wtg

Mini-batch SGD \rightarrow See K samples - update

online learning \rightarrow See 1 sample update.

Adv:

$$\left(\frac{Jb}{wb} \right) J - W \rightarrow W$$

learn and adapt quickly (1 sample very fast update)

Disadv:

learning becomes noisy and lg

task, eg: one sample says move right, other says move left, other say left not balanced.