Quiz:

- Explain the difference between Batch, Mini-Batch, and stochastic gradient descent?
- Gradient descent (GD) can always achieve the global minimum. (True/False)
- To prevent GD from overshooting the minimum:
    a) Increase the learning rate.
    b) Decrease the learning rate.
    c) Learning rate does not have any effect.
- In flat regions of the loss function gradient is:
    a) Small.
    b) Large.
    c) No relation.
- In steep regions of the loss function gradient is:
    a) Small.
    b) Large.
    c) No relation.
- Momentum-based GD (select all correct statements):
    a) Find the minimum faster than vanilla GD.
    b) Doesn’t fluctuate around the minimum.
    c) Has a lot of U-turns near the minimum.
- NAG overcomes the momentum-based GD drawbacks by the looking ahead step. (True/False)