September 2020

Sudha Devi
Roll No: 2020900003

1 Objective Question [1 mark]

Select the true statements from below:

1. The computational cost of batch gradient descent is higher than that of stochastic gradient descent as batch gradient descent requires summing over all examples.

2. The computational cost of stochastic gradient descent is higher than that of mini-batch gradient descent as the mini-batch gradient descent requires summing over a few examples.

3. In cases where multiple local minima with respect to the objective function $F(\theta)$ are present, stochastic gradient descent can sometimes avoid falling into these local minima because it uses various $\nabla_d F(\theta)$ rather than $\nabla F(\theta)$. (Here $\nabla_d F(\theta)$ are the gradient defined for each individual training example d.)

2. True. Yes computational cost of stochastic gradient descent is higher than mini batch gradient descent due to fewer updates

3. Yes, stochastic gradient descent avoids getting stuck at local minima.