

**Name :** Ankit Dadarwala

**ID :** 30

**Batch No:** 1832

**Internship**19

**Subject :** Machine learning Worksheet 1

1. B.
2. C.
3. B.
4. C.
5. A.
6. B.
7. C.
8. C.
9. A,B
- 10.B,C
- 11.C,D

12. You could use batch gradient descent, stochastic gradient descent, or mini-batch gradient descent. SGD and MBGD would work the best because neither of them need to load the entire dataset into memory in order to take 1 step of gradient descent. Batch would be ok with the caveat that you have enough memory to load all the data.

The normal equations method would not be a good choice because it is computationally inefficient. The main cause of the computational complexity comes from inverse operation on an  $(n \times n)$  matrix.

13. The Gradient descent suffer from features of different scale because the model will take longer time to reach to global maximum

We can scale the features to eliminate this problems using Standard Scaler as well