## **Summary**

Which algorithm do we use in practice?

1. **Algorithms**:
   1. GD
   2. Momentum based GD
   3. Nesterov Accelerated GD
   4. AdaGrad
   5. RMSProp
   6. Adam
2. **Strategies**:
   1. Batch
   2. Mini-Batch (32, 64, 128)
   3. Stochastic
3. Up till this point, we have been using GD with Batch update
4. In practice, Adam with Mini-Batch is the most popular choice
5. However, GD with Stochastic update is also used along with the use of special strategies to adjust the learning rate.