Assignment 5

1. Which Linear Regression training algorithm can you use if you have a training set with millions of features?
2. Can Gradient descent get stuck in a local minimum when training a logistic regression model?
3. Do all Gradient Descent algorithms lead to the same model, provided they run for the same no. of epochs?
4. Suppose you are using a Polynomial Regression and you plot the learning curves and you notice there is a large gap between training and validation error. What is the problem? How to solve it?
5. Suppose the features in your training dataset are in different scales. Which algorithms will suffer from this? How to handle this situation?