Assignment 6

1. What are the reasons for feature scaling?
   Scaling the features **makes the flow of gradient descent smooth and helps algorithms quickly reach the minima of the cost function**. Without scaling features, the algorithm may be biased toward the feature which has values higher in magnitude.

2. What is the difference between Feature Selection under Feature Engineering? Can you perform feature selection using regularization, if yes then how?
   There are many features to a single dataset and it is important to choose the best features among them that is able to give best results, by dimensionality reduction and reducing model complexity. The features that are highly impacting the result should be chooses using **feature selection** algorithms. **Features selection** is the process of selecting most suitable features for the model building.
   Feature engineering is the process of using domain knowledge to extract new variables from raw data that make machine learning algorithms work.
   Feature selection can be done using **Lasso (L1) regularization**. L1 regularization in many models causes many parameters to equal zero, so that the parameter vector is sparse.

3. Suppose you are working on a Machine Learning problem, your training accuracy is lower than the testing accuracy, what can be the reason for this?
   The **training accuracy** is lower than **testing accuracy** in case the same source dataset in not used for test or proper **train/test split** is not done in which both of them have the same underlying distribution.

4. You are training a machine learning model, your training and the testing accuracy are decreasing, what can be the reason for this?
   Decreasing accuracy as training progresses means that the learning rate for your model is too high. The model weights are constantly changing due to the high learning rate and therefore moving away from the local minimum where accuracy would be at its highest. The learning rate should be decreased to see the performance. GridSearchCv should be used to select the best learning parameter that gives optimum result.

5. What solutions you can provide for optimal bias-variance levels in a machine learning problem?
   An optimal balance of bias and variance should never over fit or under fit the model.

- For model with high bias, **linear models should not be used**, If the relation between independent and dependent variables are not linear.
- The correct way to tackle high variance will be to train the data **using multiple models**.
- **Ensemble learning methods** also help to leverage both weak and strong learners in the model to improve the model prediction.
- Another way can be to ensure that the **training data is diverse and represents all possible groups or outcomes.**
- Ensemble technique is one of the best solution in this area.