**How to select the value of K in the K-NN Algorithm?**

•There is no particular way to determine the best value for "K", so we need to try some values to find the best out of them. The most preferred value for K is 5.

•A very low value for K such as K=1 or K=2, can be noisy and lead to the effects of outliers in the model.

•Large values for K are good, but it may find some difficulties.

•We notice that based on the k value, the final result tends to change. Then how can we figure out the optimum value of k? Let us decide it based on the error calculation for our train and validation set (after all, minimizing the error is our final goal!).

•Have a look at the below graphs for training error and validation error for different values of k.

•For a very low value of k (suppose k=1), the model overfits on the training data, which leads to a high error rate on the validation set. On the other hand, for a high value of k, the model performs poorly on both train and validation set. If you observe closely, the validation error curve reaches a minima at a value of k = 9. This value of k is the optimum value of the model (it will vary for different datasets). This curve is known as an ‘elbow curve‘ (because it has a shape like an elbow) and is usually used to determine the k value.

•You can also use the grid search technique to find the best k value.

**Advantages of KNN Algorithm:**

•It is simple to implement.

•It is robust to the noisy training data

•It can be more effective if the training data is large.

**Disadvantages of KNN Algorithm:**

•Always needs to determine the value of K which may be complex some time.

•The computation cost is high because of calculating the distance between the data points for all the training samples.