

1. How does K-Nearest Neighbors algorithm work? Please answer with detailed explanation more than what I had told in the class!
 - Load the desired data.
 - Choose the value of k.
 - For getting the class which is to be predicted, repeat starting from 1 to the total number of training points we have.
 - The next step is to calculate the distance between the data point whose class is to be predicted and all the training data points. Euclidean distance can be used here.
 - Arrange the distances in non-decreasing order.
 - Assume the positive value of k and filtering k lowest values from the sorted list.
 - We have top k top distances.
 - Let k_a represent the points that belong to the a th class among k points.
 - If $k_a > k_b$ then put x in the class

2. What are the advantages and disadvantages of K-Nearest Neighbors algorithm?

Advantage	Disadvantage
A simple algorithm that is easy to understand.	The requirement of high storage.
Used for nonlinear data.	Prediction rate slow.
The versatile algorithm used for both classification as well as regression.	Stores all the training data.
Gives high accuracy but there are more good algorithms in supervised models.	The algorithm get slower when the number of examples, predictors or independent variables increases.
The algorithm doesn't demand to build a model, tune several model parameters, or make additional assumptions	

3. What are the advantages and disadvantages of Logistic Regression and when does it applied?

Advantage	Disadvantage
Logistic regression is easier to implement, interpret, and very efficient to train.	If the number of observations is lesser than the number of features, Logistic Regression should not be used, otherwise, it may lead to overfitting.
It makes no assumptions about distributions of classes in feature space.	It constructs linear boundaries.
It can easily extend to multiple classes(multinomial regression) and a natural probabilistic view of class predictions.	Logistic Regression requires average or no multicollinearity between independent variables.
It can interpret model coefficients as indicators of feature importance.	Non-linear problems can't be solved with logistic regression because it has a linear decision surface. Linearly separable data is rarely found in real-world scenarios.
Good accuracy for many simple data sets and it performs well when the dataset is linearly separable. It is very fast at classifying unknown records.	

4. Explain the differences between Decision Tree and Random Forest classifier!

Decission Tree	Random Forest
There is possibility of overfitting.	Reduce risk of overfitting.
Gives less accurate result.	Gives more accurate result.
Simpler and easier to understand, interpret and visualize.	Comparatively more complex.
A decision support tool that use a tree like grab or model of decision their possible consequences, including chance event outcomes, research cost, and utility.	An assemble learning method that of rage by constructing a multitude of decision trees and training time and output in the class depending on The individual trees.

5. List points from every aspect of differences among ensemble methods!

Bagging	Boosting	Stacking
Random	Giving mis-classified samples higher preference	Various
Minimize Variance	Increase predictive force	Both
Random subspace	Gradient descent	Blending
(Weighted) Average	Weighted majority vote	Logistic Regression

6. Does Support Vector Machine need feature scaling before the data is trained? Why? Explain it!

Answer : Yes it is. Because svm is a technique for make predictions, both predictions in case regression as well as the classification used to get separator function (hyperplane) which is optimal for separate observations that have a value of different target variables.