1)What is an ensemble technique in machine learning?

Ans- Ensemble methods are techniques that create multiple models and then combine them to produce improved results. Ensemble methods usually produces more accurate solutions than a single model would. This has been the case in a number of machine learning competitions, where the winning solutions used ensemble methods.

2) Why are ensemble techniques used in machine learning?

Ans- Ensemble methods are techniques that create multiple models and then combine them to produce improved results. Ensemble methods usually produces more accurate solutions than a single model would. This has been the case in a number of machine learning competitions, where the winning solutions used ensemble methods.

3) What is bagging?

Ans- Bagging, also known as bootstrap aggregation, is the ensemble learning method that is commonly used to reduce variance within a noisy dataset. In bagging, a random sample of data in a training set is selected with replacement—meaning that the individual data points can be chosen more than once.

4) What is boosting?

Ans- Boosting is a method used in machine learning to reduce errors in predictive data analysis. Data scientists train machine learning software, called machine learning models, on labeled data to make guesses about unlabeled data.

5) What are the benefits of using ensemble techniques?

Ans-

Ensemble methods have higher predictive accuracy, compared to the individual models.

Ensemble methods are very useful when there is both linear and non-linear type of data in the dataset; different models can be combined to handle this type of data.

6) Are ensemble techniques always better than individual models?

Ans- There are two main reasons to use an ensemble over a single model, and they are related; they are: Performance: An ensemble can make better predictions and achieve better performance than any single contributing model. Robustness: An ensemble reduces the spread or dispersion of the predictions and model performance.

7) How is the confidence interval calculated using bootstrap?

Ans-

Start with resampling with replacement from original data n times.

For each bootstrap calculate mean x\*.

Compute δ\* = x\* − x for each bootstrap sample (x is mean of original data), sort them from smallest to biggest.

Choose δ. 1 as the 90th percentile, δ.

8) How does bootstrap work and What are the steps involved in bootstrap?

Ans- Choose a number of bootstrap samples to perform.

Choose a sample size.

For each bootstrap sample. Draw a sample with replacement with the chosen size. Calculate the statistic on the sample.

Calculate the mean of the calculated sample statistics.

9) A researcher wants to estimate the mean height of a population of trees. They measure the height of a

sample of 50 trees and obtain a mean height of 15 meters and a standard deviation of 2 meters. Use

bootstrap to estimate the 95% confidence interval for the population mean height.