1. In the sense of machine learning, what is a model? What is the best way to train a model?
   A model is output file which is trained on a particular algorithm in a specific language which is capable of take raw data and produce the required output as per the training done. For example, an ML model for computer vision might be able to identify cars and pedestrians in a real-time video. The best way to train the model is that we need to use all the hyper parameters needed so that it can take care of all overfitting and other situations which can lead to produce wrong output. The model needs to be a generalised model which can work on any new or real time data.
2. In the sense of machine learning, explain the "No Free Lunch" theorem.
   It mainly states that all the optimization techniques work well equally when their performance is averaged across all platforms.

3. Describe the K-fold cross-validation mechanism in detail.

In this case if we have 100 training records and we select 5 as K value then no of experiments recorded are divided in to 5 sets. Each experiment will have 20 as test data remaining goes for training. Recurrent experiment will have 20+20 as test data and remaining goes for training. Here we take mean of the accuracy value to come at the overall result. Only disadvantage is if the training data set is not covering all forms of data then the training does not go well, hence overall performance to predict the results will not be accurate.

4. Describe the bootstrap sampling method. What is the aim of it?

As per this technique multiple subsets of data is created from original dataset with replacement, A base model is created on all the subsets. Models are run in parallel and independent of each other, final predictions is confirmed after the we combine all the predictions of all the models.

5. What is the significance of calculating the Kappa value for a classification model? Demonstrate how to measure the Kappa value of a classification model using a sample collection of results.

It can also be used to assess the performance of a classification model. For example, if we had two bankers and we asked both to classify 100 customers in two classes for credit rating (i.e., good and bad) based on their creditworthiness, we could then measure the level of their agreement through Cohen’s kappa.

6. Describe the model ensemble method. In machine learning, what part does it play?

Ensemble methods combine several decision trees classifiers to produce better predictive performance than a single decision tree classifier. The main principle behind the ensemble model is that a group of weak learners come together to form a strong learner, thus increasing the accuracy of the model.

7. What is a descriptive model's main purpose? Give examples of real-world problems that descriptive models were used to solve.

The main purposes of descriptive models are to correctly reflect the internal data structure that allows the identification of the most important regularities and dependencies. Descriptive models include exploratory data analysis models, analysis main components, factor analysis and log-linear analysis.

8. Describe how to evaluate a linear regression model.

Linear regression can be validated by checking the Mean absolute error, Mean squared error or root mean squared error. Additionally we can also check the adjusted r squared error.

9. Distinguish :

1. Descriptive vs. predictive models

- Descriptive Analytics will give you a vision into the past and tells you: what has happened? Whereas the Predictive Analytics will recognize the future and tells you: What might happen in future?
- Descriptive Analytics uses Data Aggregation and Data Mining techniques to give you knowledge about past but Predictive Analytics uses Statistical analysis and Forecast techniques to know the future.

2. Underfitting vs. overfitting the model

Overfitting: Good performance on the training data, poor generalization to other data.

Underfitting: Poor performance on the training data and poor generalization to other data

3. Bootstrapping vs. cross-validation

It is a piece of cake to implement the validation set approach. However, there is a cost for it. Because what goes in a training and validation set is determined randomly, a test error can differ greatly depending on which observations are included in each set. Therefore, the validation estimate of a test error varies a lot.

10. Make quick notes on:

1. LOOCV.

LOOCV uses a single observation as a validation set and all the rest - which is n-1 - as a training set. Then, it repeats fitting and evaluating n times by selecting a different value as the validation set. So at the end of iterations, it has n test errors. The average of these becomes the LOOCV estimate.

2. F-measurement

The F-measure is the harmonic mean of your precision and recall. In most situations, you have a trade-off between precision and recall.

3. The width of the silhouette

The Average Silhouette Width (ASW) is a popular cluster validation index to estimate the number of clusters. The question whether it also is suitable as a general objective function to be optimized for finding a clustering is addressed.

4. Receiver operating characteristic curve

ROC analysis is useful for evaluating the performance of the diagnostic tests and generally for evaluating accuracy of statistical model.