1. What are the three stages to build the hypotheses or model in machine learning?

1st step is to do data pre-process eg:if the data is in text we need to convert using some methods

-There could be duplicate values or repetead values or null values so we need to remove it and all..so there are no of data processing techniques we will discuss..

2nd step---Testing Set generation

So we have created model like we found w,b so we need to test it whether it is helpful or not..

So we have to slit our input into training and test set to cal it..

3rs step--algorthim

2. What is the standard approach to supervised learning?

Supervised learning evaluate/infer the model using Test set eg:we predicted 1900 correct and 100 wrong out of 2000.

It is like we have the label(output predciation from the i/p) eg:stock market prediction. x--i/p y--o/p

3. What is Training set and Test set?

In Machine Learning, a training set is a dataset used to train a model. In training the model, specific features are picked out from the training set. These features are then incorporated into the model. Thereby, if the training set is labeled correctly, the model should be able to learn something from these features.

Test Set: The test set is a dataset used to measure how well the model performs at making predictions on that test set. If the prediction scores for the test set are unreasonable, we’ll have to make some adjustments to our model and try again

4. What is the general principle of an ensemble method and what is bagging and

boosting in ensemble method?

Ensemble methods aim at improving the predictive performance of a given statistical learning or model fitting technique. The general principle of ensemble methods is to construct a linear combination of some model fitting methods, instead of using a single fit of the method.

Boosting could generate a combined model with lower errors as it optimises the advantages and reduces pitfalls of the single mode.

5. How can you avoid overfitting ?

If the training error is low but the Test Error is very High then we call it as overfititng.

Overfitting is more likely with nonparametric and nonlinear models that have more flexibility when learning a target function. As such, many nonparametric Machine Learning algorithms also include parameters or techniques to limit and constrain how much detail the model learns.