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

Ans: The three stages to build the hypothesis are:

1.Model building

2.Model Testing

3.Applying the model

2.**Whats is the standard approach to supervised learning?**

Ans: The standard approach to supervised learning is to split the samples into

1.Training Data

2.Testing Data

3.**What is training set and testing set?**

Ans: **Training** **set**: The set of data which is used to discover the potential predictive relationship is known as training set or the data set on which we train our model which we can use to predict the outcomes for any given features values is training data or training set.

**Testing set:** The set of data which is used to test the accuracy of our model is known as testing set.

**4.What is the general principle of an ensemble method and what is bagging and bosting in boosting in ensemble method?**

**Ans:** The general principle of an ensemble method is to combine the predictions of several models built with a given learning algorithm in order to improve robustness over a single model.

Bagging is a method in ensemble for improving unstable estimation or classification schemes. Bagging both can reduce errors by reducing the variance term.

Boosting method are used sequentially to reduce the bias of the combined model. Boosting can reduce errors by reducing the variance term.

**5. How can you avoid overfitting?**

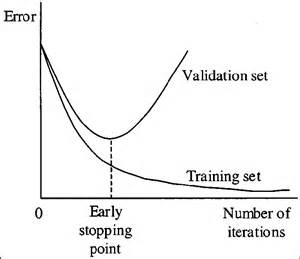
**Ans:** The most popular solution for overfitting are:

1.**Cross validation:** Use your initial train data to generate multiple mini train-test splits. Use this spits to train your model

2.**Remove features:** Remove irrelevant input features.

3.**Early stopping:** When you’re [training a learning algorithm iteratively](https://elitedatascience.com/machine-learning-iteration#model), you can measure how well each iteration of the model performs .Up until a certain number of iterations, new iterations improve the model. After that point, however, the model’s ability to generalize can weaken as it begins to overfit the training data.

Early stopping refers stopping the training process before the learner passes that point.



4.**Regularization:** Regularization refers to a broad range of techniques for artificially forcing your model to be simpler.

**5. Ensembling:** The general principle of an ensemble method is to combine the predictions of several models built with a given learning algorithm in order to improve robustness over a single model.