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

* Model building
* Model testing
* Applying the model

1. **What is the standard approach to supervised learning?**

* The standard approach to supervised learning is to split the set of examples into the Training set and the Test set

1. **What is Training set and Test set?**

* The testing dataset will only test the model while, in training dataset, the data points will come up with the model.
* A set of data is used to discover the potentially predictive relationship known as ‘Training Set’.
* Training set is an example given to the learner, while Test set is used to test the accuracy of the hypotheses generated by the learner, and it is the set of examples held back from the learner
* Training set are distinct from Test set.

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

* By using a lot of data overfitting can be avoided, overfitting happens relatively as you have a small dataset, and you try to learn from it.
* But if you have a small database and you are forced to come with a model based on that. In such situation, you can use a technique known as cross validation.
* Cross-validation is a powerful preventative measure against overfitting.

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

* 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.
* While boosting method are used sequentially to reduce the bias of the combined model.
* Boosting and Bagging both can reduce errors by reducing the variance term.

**Bagging attempts to reduce the chance overfitting complex models**.

* It trains a large number of "strong" learners in parallel.
* A strong learner is a model that's relatively unconstrained.
* Bagging then combines all the strong learners together in order to "smooth out" their predictions

**Boosting attempts to improve the predictive flexibility of simple models.**

* **I**t trains a large number of "weak" learners in sequence.
* A weak learner is a constrained model (i.e. you could limit the max depth of each decision tree).
* Each one in the sequence focuses on learning from the mistakes of the one before it.
* Boosting then combines all the weak learners into a single strong learner.