1. What are the three stages to build the hypotheses or model in machine learning?
    1) Model building
    2) Model testing
    3) Applying the model

2. 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.

3. What is Training set and Test set?

    Training Set: 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.

    Test set: 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.

4. 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.
    Boosting is a method 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?

    Steps for reducing overfitting:
    a. Add more data
    b. Use data augmentation
    c. Use architectures that generalize well
    d. Add regularization (mostly dropout, L1/L2 regularization are also possible)
    e. Reduce architecture complexity.