**Assignment 7.2**

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

**Ans:** In machine learning 3 stages involved are:

Model building : First split the data set to build model and train the model

Model Testing: second split of data to asses the performance of the model

Applying model: Applying this model on different data sets to measure the accuracy

and error to improve prediction.

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

**Ans:** The standard approach to supervised learning is to split the set of example

into the training set and the test set.

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

**Ans:** 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.

The test set is a dataset used to measure how well the model performs at making predictions on that test set.

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

**boosting in ensemble method?**

Ans: **Ensemble** is type of supervised learning model method, which combines the predictions of multiple

Smaller models to improve predictive power and generalization. Smaller models are referred

as ‘base model’.

**Bagging :** Also referred as ‘bootstrap aggregating’ in which several base estimators are built

Independently on subsets of data i.e. in parallel and their predictions are averaged

for final classifier.

**Boosting :** In this type of ensembler, base estimator are built sequentially. Each subsequent

Estimator focuses on the weakness of previous estimator by adding weight samples. Classify

Final example by majority voting.

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

**Ans:** In machine learning, when a statistical model describes random error or noise instead of underlying relationship ‘overfitting’ occurs.

To avoid overfitting, Pruning technique is used. There two types of pruning.

Pre-pruning : Stops splitting, if goodness measure is below threshold

Post-pruning: Goes deeper to build complex tree and then pruning is done.