Machine Learning Interview Questions Part 4

1. Write a basic Machine Learning program to check the accuracy of a model, by importing any dataset using any classifier?

```
#importing dataset
import sklearn
from sklearn import datasets
iris = datasets.load_iris()
X = iris.data
Y = iris.target

#splitting the dataset
from sklearn.cross_validation import train_test_split
X_train, Y_train, X_test, Y_test = train_test_split(X,Y, test_size = 0.5)

#Selecting Classifier
my_classifier = tree.DecisionTreeClassifier()
My_classifier.fit(X_train, Y_train)
predictions = my_classifier(X_test)
#check accuracy
From sklear.metrics import accuracy_score
print accuracy_score(y_test, predictions)
```

2. What are the different types of Machine Learning?

There are three ways in which machines learn:

1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

Supervised Learning:

Supervised learning is a method in which the machine learns using labeled data.

- It is like learning under the guidance of a teacher
- Training dataset is like a teacher which is used to train the machine
- Model is trained on a pre-defined dataset before it starts making decisions when given new data

Unsupervised Learning:

Unsupervised learning is a method in which the machine is trained on unlabelled data or without any guidance

- It is like learning without a teacher.
- Model learns through observation & finds structures in data.
- Model is given a dataset and is left to automatically find patterns and relationships in that dataset by creating clusters.

Reinforcement Learning:

Reinforcement learning involves an agent that interacts with its environment by producing actions & discovers errors or rewards.

- It is like being stuck in an isolated island, where you must explore the environment and learn how to live and adapt to the living conditions on your own.
- Model learns through the hit and trial method
- It learns on the basis of reward or penalty given for every action it performs

3. Explain Classification and Regression

Classification	Regression
• Classification is the task of predicting a discrete class label	• Regression is the task of predicting a continuous quantity
• In a classification problem data is labelled into one of two or more classes	• A regression problem requires the prediction of a quantity
• A classification problem with two classes is called binary, more than two classes is called a multi-class classification	• A regression problem with multiple input variables is called a multivariate regression problem
• Classifying an email as spam or non-spam is an example of a classification problem	• Predicting the price of a stock over a period of time is a regression problem

4. Explain false negative, false positive, true negative and true positive with a simple example.

Let's consider a scenario of a fire emergency:

- True Positive: If the alarm goes on in case of a fire.

Fire is positive and prediction made by the system is true.

- False Positive: If the alarm goes on, and there is no fire.

System predicted fire to be positive which is a wrong prediction, hence the prediction is false.

- False Negative: If the alarm does not ring but there was a fire.

System predicted fire to be negative which was false since there was fire.

- True Negative: If the alarm does not ring and there was no fire.

The fire is negative and this prediction was true.

5. What is a Confusion Matrix?

A confusion matrix or an error matrix is a table which is used for summarizing the performance of a classification algorithm.

n=165	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

Consider the above table where:

- TN = True Negative
- TP = True Positive
- FN = False Negative
- FP = False Positive