



---

# MACHINE LEARNING BEGINNER INTERVIEW Q&A

---

By [QuickTechie.com](http://QuickTechie.com)



**Question-1: In Common term what do you mean by Advanced Analytics?**

**Answer:** From 10,000 feed view we can say that Advanced Analytics is a collection of techniques which you are using to solve the core problem of deriving insights and making predictions or recommendations based on the data.

**Question-2: Can you give some examples which you perform as part of Machine Learning?**

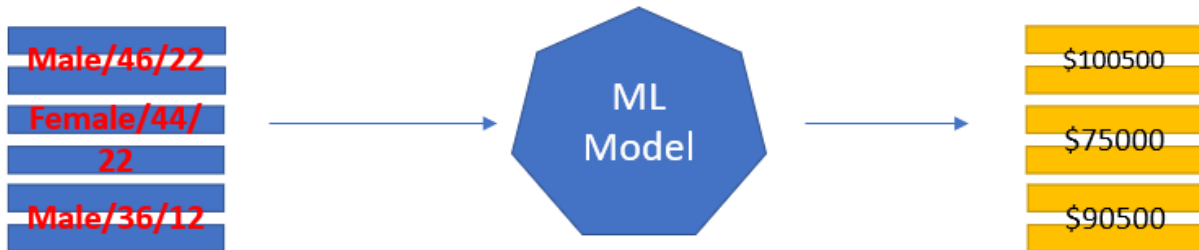
**Answer:** There 1000 different types of tasks you can perform using Machine Learning but below are the few most common task which you usually performs.

- **Supervised Learning:** Based on the data characteristics or features you try to predict a label for each given data point. And that is also part of your supervised learning.
- **Recommendations:** recommending users new products based on his or her past behavior.
- **Unsupervised Learning:** You use various different techniques to discover the structure in the data.
- **Graph Analytics:** This quite a different and wide domain in itself. Which is used to find or search the patterns in social networking.

**Question-3: Can you explain what is a supervised learning?**

**Answer:** - Supervised Learning is most commonly used Machine Learning. In this you have already have historical data with the labels and features. You would create a Machine Learning model such that using features of new data you predict the labels. For example, lets assume you have to predict salary for each individual based on his/her characteristics.

Here, label would be salary and features are his or her age, years of experience, gender etc. You already have some historical data with this information and you need to predict salary for new data.



**Salary :** Dependent Variable. (Label :To be predicted)

**Features:** Age, Experience, Gender

©HadoopExam.com

**Question-4: What do you mean by dependent variable in the Supervised Machine Learning?**

**Answer:** When you run Machine Learning Models for supervised learning and expect some output as part of the Machine Learning model completion which is usually some predictions. Like based on the number of years of experience, Age and Gender your model predict salary. This predicted label is called dependent variable. Conceptually, you can think of this label (salary) is depend on the features.

**Question-5: What are the independent variables in supervised learning?**

**Answer:** In Supervised Machine Learning you need to predict (dependent variable) based on the independent variable like age, gender, experience etc. And these independent variables are also referred as **Features**. Independent variables are the input to your Machine Learning model which would be analyzed to generate output (independent variables or labels).

**Question-6: Can you please give another example of independent variables and dependent variables?**

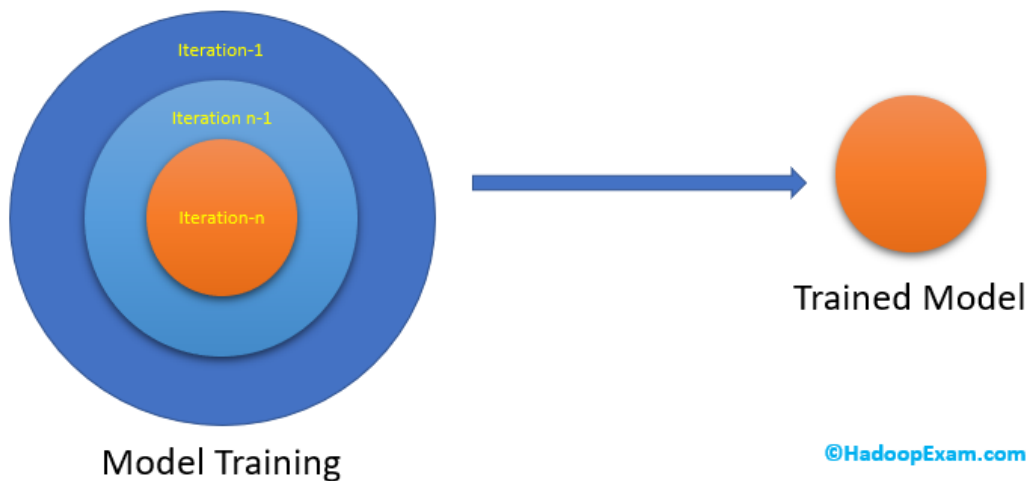
**Answer:** Yes, you would have all the click data from the website pages and user demographics. Like his age, education, duration on web page, pregnant etc. And based on that you want to predict whether product was purchased or not on the website.

**Question-7: Is it true that usually Machine Learning algorithm can generate results in single iterations?**

**Answer:** No, this is not true. If that have been a case, then no new technologies were needed to support and run Machine Learning algorithms. Almost, every Machine Learning algorithm required iterative processing on the data to generate or predict final and more accurate dependent variable.

**Question-8: Can you give a reason why multiple iterations are needed in Machine Learning Model training?**

**Answer:** In Machine Learning Model required iterative runs, because in each run it optimizes the algorithms. Your training model will start with a basic model and improves with each iteration by adjusting various internal parameters such as coefficients in each training iterations. And once all the iteration is finished, you would have a model which is well trained to make predictions on new data. Same happens with human being as well, as much as he gains experience in his or her life, he can well predict the planning for new project, that's why a well-trained human being is paid more as well.



**Question-9: In common words, what you have to do after model training before using it?**

**Answer:** There are various tasks which we need to complete before you use your trained model like below.

- Measure the success of your trained models.

Once, you are satisfied with your model. You put them in the production to predict the new values.

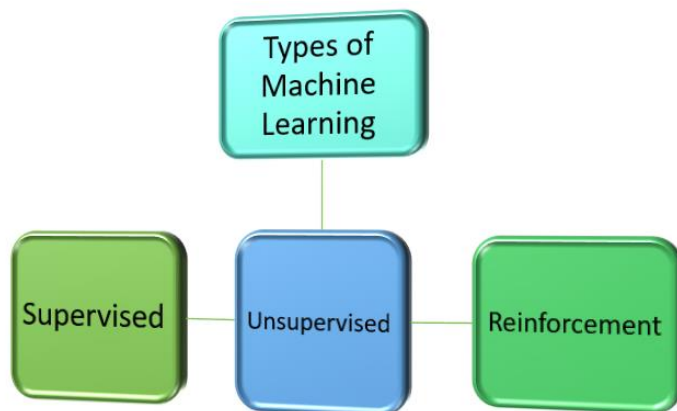
**Question-10: Can you define Machine learning once again in different words?**

**Answer:** Yes, we can say that Machine Learning is to create a Model using computer program or algorithm by teaching it again and again until it become accurate or very near to accurate for a particular task which you want to accomplish.

**Question-11: Which are major categories in which you can divide Machine Learning?**

**Answer:** Machine Learning can be classified in three categories as below.

- Supervised Machine Learning.
- Unsupervised Machine Learning.
- Reinforcement Machine Learning.



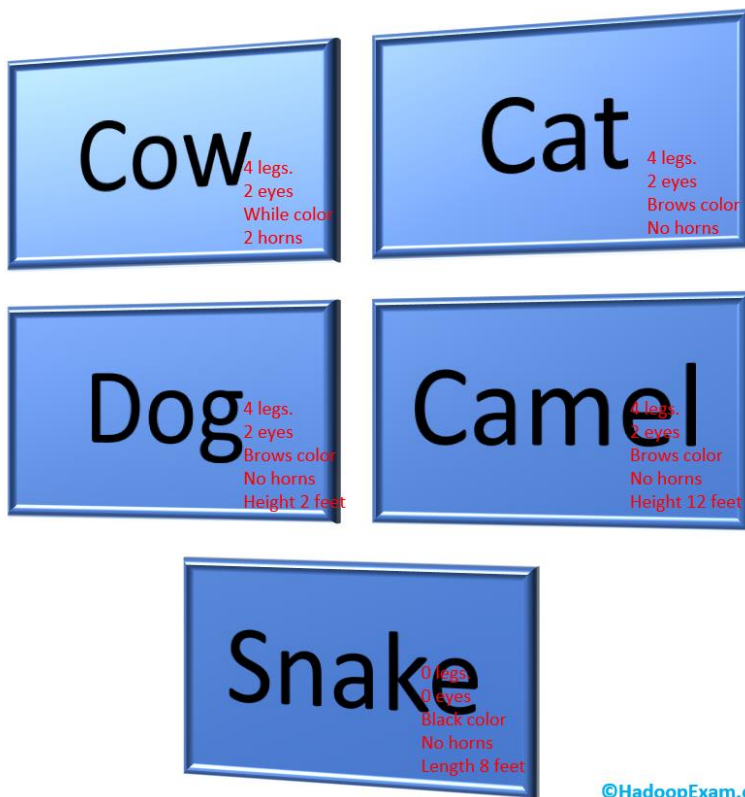
©HadoopExam.com

**Question-12: What do you mean by reinforcement learning in Layman words?**

**Answer:** In Layman words you can say, Reinforcement Learning is “Learn by Mistakes”.

**Question-13: Which type of Machine Learning is relative easier to implement among all three?**

**Answer:** Supervised Machine Learning is relatively easier to implement as well as understand. You can think like you are teaching very young children with the flashcards. Flash cards has labels (dependent variables) and some features in the form of picture (that is independent variables).



**Question-14: Again, why multiple iteration during Machine Learning Model Training?**

**Answer:** When you train your machine Learning Model, you keep feeding new data and check it has predicted correct label or not. If not, you adjust your algorithms, and again feed the data and check its predicted value. Over time, the algorithm would learn to approximate the exact nature of relationship between input labels and its features (dependent variable) or characteristics (independent variables).

**Question-15: What happened when your Supervised Machine Learning Algorithm is fully trained?**

**Answer:** Once, your Machine Learning algorithm is fully trained then you can productionize it and would be able to observe new and never before seen data and predict almost correct label for that new data.

Known Input

Male/46/22  
Female/44/  
22  
Male/36/12



Known Output

\$100500  
\$75000  
\$90500

**Salary** : Dependent Variable. (Label :To be predicted)

**Features**: Age, Experience, Gender

©HadoopExam.com

Trained Model in Production

Male/50/28  
New Data



\$110500

Predicted Value

©HadoopExam.com

**Question-16: Can you give some more detail about supervised learning?**

**Answer:** We can describe supervised learning as task-oriented and which is highly focused on the singular tasks. And you should have as much as possible example data to train your model, so that in production it can predict dependent variable more accurately.

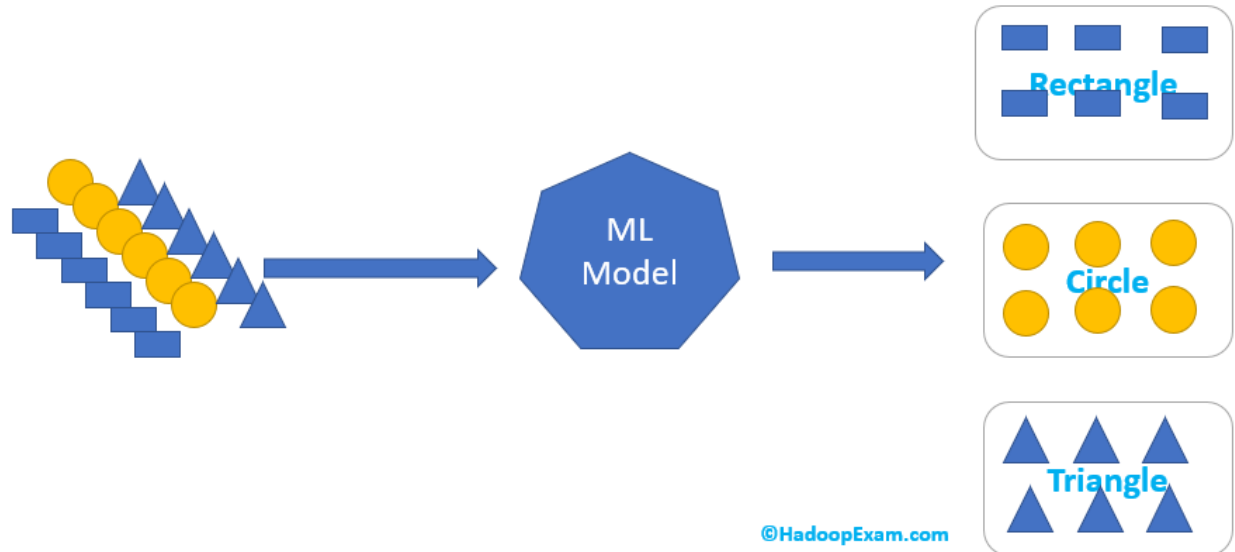
**Question-17: Can you give few examples of Supervised Learning?**

**Answer:** Yes, below are the few examples

- **Gmail Auto classifications:** You have seen in your Gmail box Google classify the emails whether an email is for marketing, social networking, spam or most important etc.
- **Spam filter:** Your new email based on contents can be classified whether it's a spam or not. As well also check your response as well to each email. Yes, email provider reads your email to correctly classify it.
- **Face recognition:** Now in smart phone, you have seen that Camera automatically capture the face. So, your mobile has Supervised Machine Learning trained model in it and while you capture photos, they can suggest faces etc.

**Question-18: What is un-supervised learning?**

**Answer:** You can think of its an opposite of supervised learning. It does not have pre-defined labels. (There is nothing which should supervise the Model). But you would feed whatever data you have to Machine Learning Model and based on the features and characteristics it generates the groups or you can say cluster. Which you can easily understand the way data is organized or grouped by Model.



**Question-19: Why unsupervised learning is more challenging?**

**Answer:** If you see most of the data currently is unlabeled (there are various jobs opening in recent years to just label the data, so that supervised learning can be used). And it is very interesting to have such an intelligent Machine Learning model which can accurately groups the Tera-bytes of unlabeled data. And make sense out of these unlabeled data is very demanding because these data were lying idle and previously was of no use and was discarded. But now that is not the case, every company want to make intelligent decision based on this idle unlabeled data as well. And day by day new unlabeled data is being generated.

**Question-20: Can you give an example of unsupervised learning?**

**Answer:** Yes, lets assume you have 1 Million research papers on various topics. Now you want to group all research papers in respective research areas. However, its not known in advanced what all research areas these papers are written. Hence, you will be running one of the un-supervised Machine Learning algorithm which can group all these papers in respective or related research areas.

**Question-21: Why do you say, unsupervised Machine Learning is data driven?**

**Answer:** Because unsupervised learning is based upon the data and characteristics/features/properties of the data. And also, the outcome from an unsupervised learning tasks are controlled by these data. Hence, it is called data driven Machine Learning.

**Question-22: Is recommender system is part of unsupervised learning or supervised learning?**

**Answer:** This would fall into the un-supervised learning. Let's take an example of online video platform like YouTube. On YouTube following characteristics are known in advance.

- Title of Video.
- Length of video.
- Genre of the video.

And history of the videos watched by a user. Taking all these data into account, YouTube needs to recommend or suggest new videos that you might have not seen yet. So recommender system will try to find the relationship in the data and based on that it will suggest you new video to watch.

**Question-23: Can you give couple of more example for un-supervised learning?**

**Answer:** Below are the examples of unsupervised learning

- Grouping customer based on buying habits.
- Grouping user complains.

**Question-24: In one line can you tell me the difference between supervised and un-supervised learning?**

**Answer:** Yes, Supervised learning has the label already known to predict. But Un-supervised does not know in advance. And these labels are also known as dependent variables.

**Question-25: Email Spam detection is what kind of classifications?**

**Answer:** You, can consider email spam detection is as a Binary Classification under supervised learning. Because, here you are trying to find that particular email is a spam or not. Only one of the two possible labels.

**Question-26: Can you tell me; how would you productionize the Spam Detection Machine Learning Algorithm?**

**Answer:** Let's say, we have 100,000 of historical emails, which are already labeled like Spam or not Spam. We define our Machine Learning algorithm and train it with these all historical emails, until it can predict all the emails correctly. Once we get almost 100% accuracy, we can put this Machine Learning Model or algorithm in production to categorize all new future emails, whether those are spam or not spam.

**Question-27: What is multi-class classification?**

**Answer:** As name suggests, when we classify items in more than two categories, that is called Multiclass Classification.

**Question-28: Can you few more examples of classification Machine Learning?**

**Answer:** Yes, below are few more examples to clear our understanding about classifications.

- **Disease Prediction:** Based on historical information from millions of patient's responses a model can be created and trained to predict the disease for the new patients. We are also seeing now a day's online applications are available they ask few questions and based on that they predict possible disease.
- **Customer Churn:** When customer can stop using services, like unsubscribing from a subscription. We can use all our existing and past customer. Based on their behavior we can



create a binary classification model. Whether or not the subscriber would continue our services or not.

- **Prediction on Buy or not Buy online:** You are having an eCommerce website and need to find whether customer will buy a product or not on the website. To do that, you can use previous history of the data and train your model accordingly for instance browsing patterns, customer location etc. And your model can predict whether visitor is a potential customer or not. And based on that you can pop some offers for him instantly.

**Question-29: What is the regression technique in Machine Learning?**

**Answer:** As with the classification Machine Learning technique we tried to predict some discrete values like yes or no. Buy or not buy. Spam or not spam etc. However, in regression we would have to predict some continuous variables also you can say real number. So rather than you predict category, your model will be predicting number on some line. However, all other stuff is same as other classification Machine Learning model. Hence, we can say regression is also a Supervised Machine Learning.

**Question-30: Can you give some examples of regression (Supervised Machine Learning)?**

**Answer:** Yes, below are the few examples of regressions

- **Online Sales Prediction:** Use last 10 years sales data and create and train model using this data to predict the next month sales.
- **Height Prediction:** Based on the historical data which include Parent's height, location e.g. Hilly area, forest area, plain area etc. You want to predict the height of the children based on their age and what maximum possible height they can reach.
- **Prediction of viewers on day one:** You want to have a model based on the character's in the movie, story of the movie, current trending topic, epidemic etc. And model should predict the number of possible viewers on day1, day2, day3 of the movie release.

Question-31: Which all Machine Learning Model you think can be part of the Supervised Learning?

**Answer:** We can say at least following 3 falls under the supervised Learning.

- **Classification:** (This can be supervised as well as unsupervised), based on that whether we already have the labels to classify or not.
- **Regression:**
- **Recommendation system:** (This can be supervised as well as unsupervised), based on that whether we already have the labels to predict.

**Question-32: Which Supervised Machine Learning is best suited in Apache Spark or Big data world?**

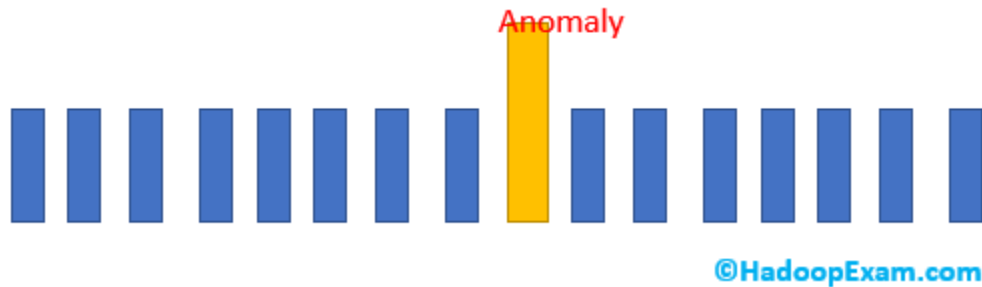
**Answer:** Recommender System is a very common use case with the Spark and also well suited under the Big Data.

**Question-33: can you once again tell me something about un-supervised learning?**

**Answer:** Yes, using unsupervised Machine Learning we are trying to find the patterns or try to discover the underlying structure in the given data. Without having any dependent variable or prediction variables in advance.

**Question-34: What do you mean by Anomaly detection?**

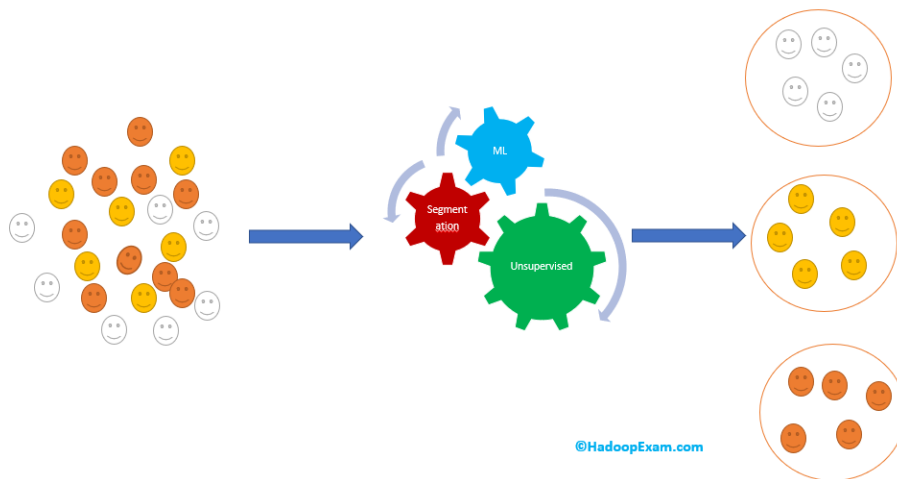
**Answer:** Suppose you have some regular events happening on time or on regular interval. However, you want to catch any unusual activity during this time or interval which had not happened in past or have very different from other events. You can use Anomaly Detection algorithm to catch it. Like in share market millions of transactions are happening on daily basis. But you want to check for some unusual transaction to detect the fraud.



Another example is traffic police control want to detect any un-usual activity on the common road or path.

**Question-35: Can you give an example of user segmentations?**

**Answer:** A blogging website, want to recommend new blogs based on the user behavior, time spent on the website etc. They can put each user in different possible groups and based on the time spent or study or browsing they do.

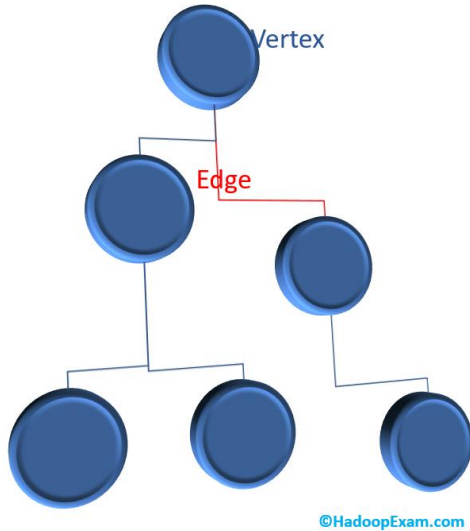


**Question-36: What do you mean by Topic Modeling?**

**Answer:** This is also an example of un-supervised learning. Like you want to group all the research papers based on their respective area, which are not known in advance. You use Topic modeling, in which you find the related words in each individual document and then group them together. So that they can be segmented on devised group once they are processed by your Model.

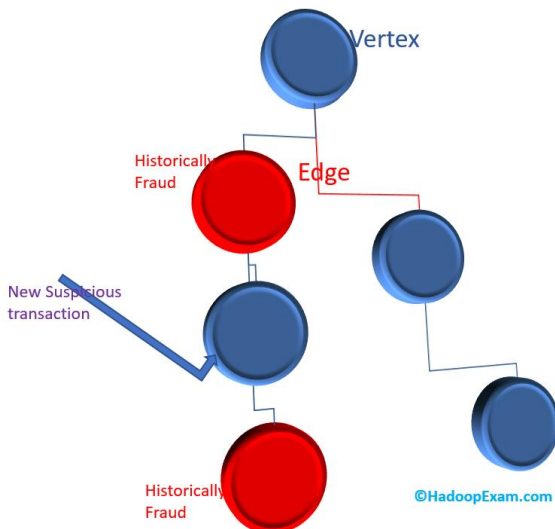
**Question-37: In Graph Analytics, what vertices and edges represent?**

**Answer:** In graph analytics vertices represents the objects and edges represent relationships among the objects. For example, Objects can be human being in the social networking and edges represent relationships between two human beings. Once you take a look of properties for the vertices and edges, you can better understand the relationship between objects.



**Question-38: Can you use graph analytics for fraud detection between transactions?**

**Answer:** Yes, we can. Let's assume we already have historical data like phone numbers, addresses, email ids etc. from historical fraud transactions. Now you want to catch any new fraud or suspicious transactions. You can build a model which can capture the transactions between two fraud phone numbers or contacts. Which can be possibly a fraud transaction on newer way.



**Question-39: Can you give another example; how can you use say that a particular relationship is suspicious?**

**Answer:** Lets assume you have a graph with the created with your data. And on average you found that each individual vertex has 8-10 edges. But there are 2 nodes/vertex in entire graph has only 1 edge. Then you can consider that relationship as a suspect.

**Question-40: Can we use to determine the influencer in social graph using graph analytics?**

**Answer:** Yes, it is possible. Let's assume you have social graph created and you can classify individual as an influencer based on its edges count or other properties. And this can be used as a classifier to check an individual an influencer or not. If individual is an influencer then companies can use that influencer to advertise their brands.

**Question-41: Can graph analytics can be used for recommendations?**

**Answer:** Yes, it can be used. Let's say, you are developing a search engine for your internal website in organization. And you want to arrange the pages based on their importance. You can create a graph between the page links. And the page which has highest link coming to it from other different pages, can be consider the most important page and can be recommended on top as part of search results.

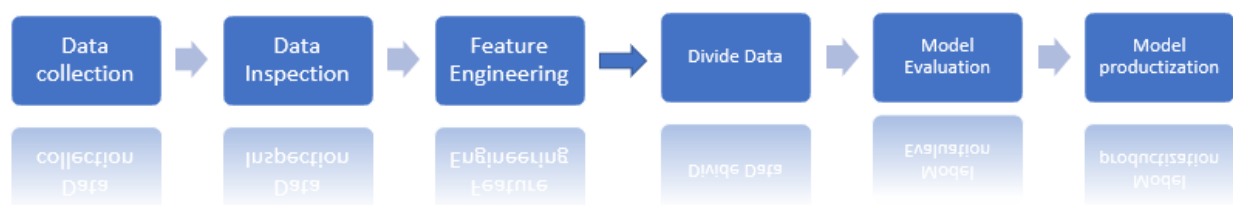
**Question-42: In general, what all steps, you have to do as part of Machine Learning advanced analytics?**

**Answer:** When you have to do advanced analytics using Machine Learning, you have to perform below tasks at least. In future question, you would get in more detail.

- Preparing your data (Data collections and format the data).
- Select Machine Learning Model.
- Testing Machine Learning Model in various different ways.
- Model Tuning.
- Productionizing the Model for newer of future data.

**Question-43: Can you again tell me in detail the steps you have take in your previous advanced analytics projects?**

**Answer:** For any new project I have to take following steps to complete any Machine Learning Advanced analytics.



[@HadoopExam.com](https://www.hadoopexam.com)

1. **Data Collection:** Collect and find all the available historical data. You may already have it in your organization data lake.
2. **Data Inspection:** You need to clean this data as per your need. Also inspect the data, whether it fulfills our need or not.

3. **Feature engineering:** Now you need to convert your data in proper format which can be used by your algorithm or model. Like converting text data into numerical vectors.
4. **Divide data:** You need to divide your available data in two parts at least. First would be used to train your model. And rest of the data can be used for testing your model. And based on the testing you can find what all candidate models you want to consider. For example you divide 60% your data for testing and 40% data for model evaluation.
5. **Model evaluation:** Now you have defined the success criteria for your model. And you would evaluate each candidate model based on your success criteria. To measure the success criteria you would use for your model. You should use 40% data from your previous step which you kept aside to evaluate your model.
6. **Productionize model:** Now you have selected your correct model. You need to put them on actual work to work on new future data.

However, please note that, you would not require all these steps for every advanced analytics. Its all depend on what data you have and what stage or format your data is available.

#### **Question-44: What do you mean by data collection in Advanced Analytics?**

**Answer:** This is one of the first step in your data analytics and also hardest steps compare to any other step. If you are looking for overall end to end project delivery. You need to gather the data from various places and sometime you need to buy from third party vendor if you don't have this data. And this data you would use to test and evaluate your Machine Learning Model. Popular tool which are used for data collection specially in Big Data world are below

- **Apache Spark:** It can help you to process and collect data from various sources and store them in one central place like HDFS.
- **Hadoop framework:** Hadoop a has various different kind of components which can help to bring the data and save them into the HDFS or cloud buckets.

#### **Question-45: What do you mean by Data Cleaning step?**

**Answer:** Once you have purchased or gathered the data which you want to use for your model building. The next step is to clean the data. And this step ins statistics is known as Exploratory Data Analysis or EDA. You should be able to run adhoc query on your data and also use some visualizations tool. And you try to understand the relationships in the data with the help of data distributions and co-relations in the data. And you want to remove some of the data which you don't need or you want to fill in some missing data etc. And this one of the critical steps to understand your data in detail. You may want to spend more time at this step to avoid mistakes in further steps or repeat this step again. If you are using Apache Spark then using [Spark SQL](#) for this.

#### **Question-46: What do you mean by Feature Engineering?**

**Answer:** This is the step where you need to convert your data in the form where it can be understood by Machine Learning algorithms. Like converting data features into numerical features. Again, this needs to be done with carefulness. Feature engineering include following steps

- Data Normalization.
- Adding variables to represent the interactions of other variables.

- Manipulating categorical variables.
- Changing into the format which can be accepted by Machine Learning Model.

And to manipulate your data, you have to use various Machine Learning statistical techniques.

**Question-47: Apache Spark MLlib Machine Learning Library, in what format input should be provided?**

**Answer:** In case of Apache Spark MLlib or Machine Learning Library, all variables usually has to be input as vectors of doubles. Hence, you have to convert your data features accordingly.

**Question-48: What all you do while Model Training step?**

**Answer:** During the model training parameters inside the model will change according to how well the model perform on the input data. Suppose you are building a Spam classification model, then the algorithm which you will be using likely to find that certain words are better predictors of spam than others and therefore weight the parameters associated with those words higher. And once your model is trained, we find that certain words should have more influence because their consistency associated with spam emails then others.

And output of this step we would have a Machine Learning model. This model you would use for future predictions.

**Question-49: How do you know that the model we trained is a good or not?**

**Answer:** For that you use Model Tuning and Model evaluation step.

**Question-50: What is the general reason because of which Machine Learning Model does perform poorly?**

**Answer:** The cause of poor performance of Machine Learning Algorithm is because of either Overfitting or underfitting of the data.

**Question-51: In Machine Learning, why generalization is important?**

**Answer:** In Machine Learning Generalization is important because the data we collected to build and test our model is only sample data and those does not cover all the use cases or possibility in the data. And most of this data is noisy as well.

**Question-52: What is the problem, if you over train your Machine Learning Model?**

**Answer:** When you overtrain your Machine Learning Model on the training data set then certainly it would identify all the relevant information in the training data. However, it would most likely fail to make correct prediction when you input completely new data to your model. And this is the situation where you can say that model is incapable of generalizing or that it is overfitting the training data. And also, you train your model more and more, it will become more complex as well.

**Question-53: What is the problem with the over trained model?**

**Answer:** As we have seen in previous question, as you train more and more your model. Its complexity increases and, on the training, data sets its error reduces. And you think your model would perform better. However, as soon as you present new data to this model it would fail in prediction.

**Question-54: What do you mean by Machine Learning Model Generalization?**

**Answer:** Whatever, model you are creating it should be able to Generalize well. It means, you should not overtrain your Model with the training data set. Hence, to predict a good Advanced Machine Learning Model it should be capable of generalizing, and you should know when to stop training the model and it does not overfit. Best example in layman term, don't teach each and every detail to your child. Let them learn something from real world. They should be able to generalize new situations based on the learning/teaching you taught them. Otherwise, they would always depend on parent and when parent is not there. They would miserably fail to perform in new situations or conditions.

**Question-55: How data quality affects the Model Generalization?**

**Answer:** Hence, to have model well generalized, you must ensure that the data used to train model is good and reliable sample of observations. The higher the quality of the data and more representative the sample is the easier it will be for model to learn the unknown mapping.

Generalization, means you predict well on the unseen data.

**Question-56: What do you mean by training set?**

**Answer:** Whatever Machine Learning Model you create those were trained on some certain datasets. And that dataset is called training dataset. And during the model learning phase Model learn from the inherent patterns in the data and loss function or objective function is reduced through training. If the trained model works or perform same as while training it did with respect to performance then you can say your model is generalized as well.

**Question-57: What is the main objective of the Machine Learning Model tuning?**

**Answer:** The main point of Machine Learning Model tuning is to generalize it, so that it can perform better with the new data.

**Question-58: Can you help in understanding of Machine Learning Model generalization in statistical term?**

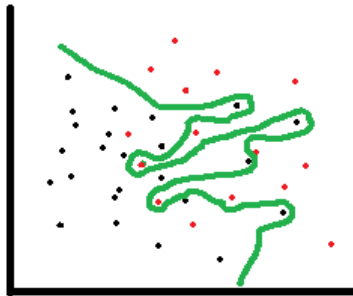
**Answer:** Generalization, means fitting your Machine Learning Model or algorithm on the sample data to eliminate the effect of Outliers, unstructured pattern, seasonality which exists in your data.

And having Generalized model do predictions on the new data which is more concrete, and having less variance in it with the minimum bias-ness.

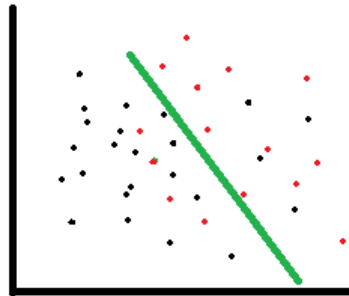
Hence, you should even apply the generalization on your training data set as well before actually you use them to train your model.

**Question-59: Can you show using graph what is Generalization and overfitting?**

**Answer:** As you can see in below image. Image in left represent a model which is over trained and had been overfitting. And would perform worst with the new real data. And you can also see this model is highly complex as well. But you see image in the right which is a generalized model and perform best when represented with the new data. Your target should always to create most generalize model, so that you can predict it very well.



**Overfitting**



**Generalized**

[HadoopExam.com](http://HadoopExam.com)

Question-60: What are your expectations from Machine Learning Model?

**Answer:** We need Machine Learning Models to describe the underline pattern in the data. However, when you work with the real data. It does not come without the noise. And we want that model capture the signal in the data and not the noise.

Question-61: Why overfitting makes model useless with the new data?

**Answer:** Overfitting happens when your model performs better with the training data and have more or high accuracy with the training data set. However, its accuracy drops when new data is given to the model. Because model is not generalized for new data and overfit with the training data set.

**Question-62: How can we avoid overfitting in Machine Learning Model?**

**Answer:** To prevent overfitting with the Machine Learning Model we can do following things

- Increase the volume of data for training: If you have more data to train your model then its accuracy increases.

There are more ways, we will discuss in later questions as you learn more concepts.

**Question-63: What do you mean by Cross Validation in Machine Learning?**

**Answer:** We use cross validation to detect model overfitting or finding whether model is failed to generalize itself. Using the cross-validation technique, we can evaluate the various Machine Learning Models using the sample test data. And later validate with other sample data to find the accuracy of the model.

**Question-64: What is the model underfitting?**

**Answer:** Your model is called underfit when it performs badly with the training data itself.

**Question-65: Why underfitting happens on the Machine Learning Model?**

**Answer:** Underfitting happens because the model you are working upon is not able to capture the relationship between the input examples and the target variables.

$$Y = f(X)$$

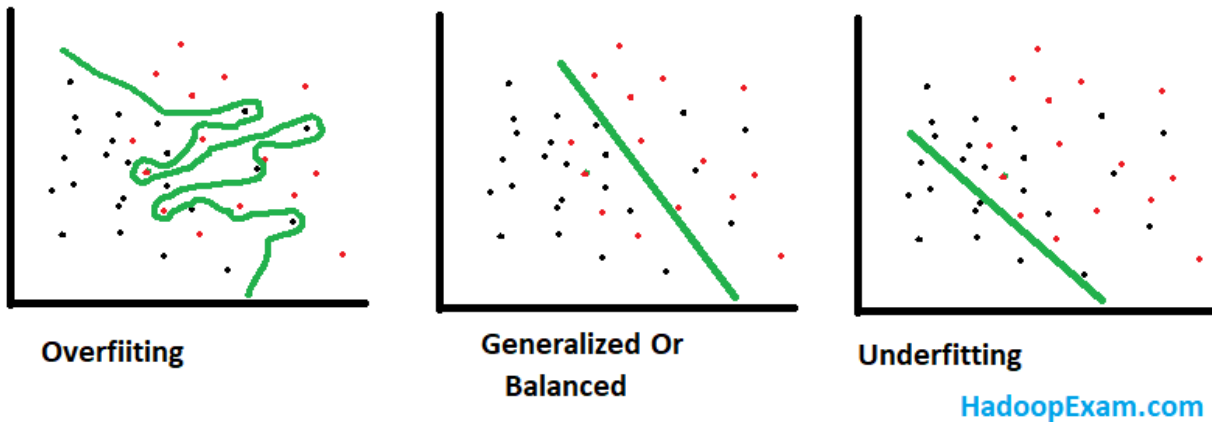


X=Input Variable.

Y= Output variable.

**Question-66: Can you graphically show the underfitting model?**

**Answer:** As you can see in below image the 3<sup>rd</sup> graph is not able to correctly predict the data with even the training dataset. Hence, it is considered underfitting. (check [HadoopExam.com](http://HadoopExam.com) machine learning training to understand further)



**Question-67: Why model overfitting happens?**

**Answer:** Because your model tries to memorize more and more data during the training data set.

**Question-68: If our model is underfitting and poorly perform what we can do?**

**Answer:** When your model perform badly with the training dataset then possible reason could be your model is very simple and your input data features are not expressive enough to describe the target output. And to improve the performance in this case can be done by increasing the flexibility of the model and to increase the model flexibility we can take following steps

- Add new features in the input training dataset.
- Reduce the regularization in the data.

**Question-69: What should you do reduce the model overfitting?**

**Answer:** To avoid overfitting with the training dataset, you can reduce the model flexibility. And to reduce the model flexibility, we can take following actions.

- Reduce the number of feature combinations.
- Reduce the number of numerical attribute bins.
- Increase the amount of regularization is used.

**Question-70: What other options do you see to increase the accuracy of the model?**

**Answer:** To increase the accuracy of the model you should not have poor data and to improve it, you can do the following things.

- Increase the amount of training data examples.
- Increase the number of iterations on the training data.

**Question-71: What is Machine Learning, can you explain once again?**

**Answer:** Using Machine Learning on the available historical data, you or your team can make better business decisions. You use different Machine Learning algorithms for different needs to find the patterns in the available data. These models are mathematical. Once the correct or near accurate model you find then you put this model in productions to make the predictions on the new or future data.

**Question-72: Can you give an example, when you don't want to use Machine Learning for predicting the future values?**

**Answer:** We should not use the Machine Learning when you can determine the future value based on some simple or existing solutions like applying some rules using rule engine or by simply doing calculations you can derive the value or you can write step by step program to derive the value.

**Question-73: Can you give the example when you really want to use Machine Learning?**

**Answer:** You don't have choice and need to go for Machine Learning only, in following situations

- When you are not able to use rules-based algorithm to derive the outcome on new data.
- When your computation can not derive the target value.
- When your computation is depending on too many factors and you can not code them in program.
- When you have to predict outcome on large volume of data. Like predicting website visitor will buy a product or not when there are millions of visitors on the website.