# **Unit 1: Introduction**

1. **What is Machine Learning? Explain the importance of ML with suitable example.**

Machine Learning is the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyses and draw inferences from patterns in data. Machine learning is an application of AI that enables systems to learn and improve from experience without being explicitly programmed. Machine learning focuses on developing computer programs that can access data and use it to learn for themselves. The machine learning process begins with observations or data, such as examples, direct experience or instruction. It looks for patterns in data so it can later make inferences based on the examples provided. The primary aim of ML is to allow computers to learn autonomously without human intervention or assistance and adjust actions accordingly.

Machine learning is important because it gives enterprises a view of trends in customer behavior and business operational patterns, as well as supports the development of new products. Many of today's leading companies, such as Facebook, Google and Uber, make machine learning a central part of their operations. Machine learning has become a significant competitive differentiator for many companies.

Machine Learning is broadly used in every industry and has a wide range of applications, especially that involves collecting, analyzing, and responding to large sets of data. The importance of Machine Learning can be understood by these important applications.

Some important applications in which machine learning is widely used are given below:

1. **Healthcare:** Machine Learning is widely used in the healthcare industry. It helps healthcare researchers to analyze data points and suggest outcomes. Natural language processing helped to give accurate insights for better results of patients. Further, machine learning has improved the treatment methods by analyzing external data on patients' conditions in terms of X-ray, Ultrasound, CT-scan, etc. NLP, medical imaging, and genetic information are key areas of machine learning that improve the diagnosis, detection, and prediction system in the healthcare sector.
2. **Automation:** This is one of the significant applications of machine learning that helps to make the system automated. It helps machines to perform repetitive tasks without human intervention. As a machine learning engineer and data scientist, you have the responsibilities to solve any given task multiple times with no errors. However, this is not practically possible for humans. Hence machine learning has developed various models to automate the process, having the capability of performing iterative tasks in lesser time.
3. **Banking and Finance:** Machine Learning is a subset of AI that uses statistical models to make accurate predictions. In the banking and finance sector, machine learning helped in many ways, such as fraud detection, portfolio management, risk management, chatbots, document analysis, high-frequency trading, mortgage underwriting, AML detection, anomaly detection, risk credit score detection, KYC processing, etc. Hence, machine learning is widely applied in the banking and finance sector to reduce error as well as time.
4. **Transportation and Traffic Prediction:** This is one of the most common applications of Machine Learning that is widely used by all individuals in their daily routine. It helps to ensure highly secured routes, generate accurate ETAs, predict vehicle breakdown, Driving Prescriptive Analytics, etc. Although machine learning has solved transportation problems, it still requires more improvement. Statistical machine learning algorithms helps to build a smart transportation system. Further, deep Learning explored the complex interactions of roads, highways, traffic, environmental elements, crashes, etc. Hence, machine learning technology has improved daily traffic management as well as a collection of traffic data to predict insights of routes and traffic.
5. **Image Recognition:** It is one of the most common applications of machine learning which is used to detect the image over the internet. Further, various social media sites such as Facebook uses image recognition for tagging the images to your Facebook friends with its feature named auto friend tagging suggestion.  
   Further, now a day's, almost all mobile devices come with exciting face detection features. Using this feature, you can secure your mobile data with face unlocking, so if anyone tries to access your mobile device, they cannot open without face recognition.
6. **Speech Recognition:** Speech recognition is one of the biggest achievements of machine learning applications. It enables users to search content without writing text or, in other words, 'search by voice'. It can search content/products on YouTube, Google, Amazon, etc. platforms by your voice. This technology is referred to as speech recognition. It is a process of converting voice instructions into the text; hence it is also known as 'Speech to text' or 'Computer speech recognition. Some important examples of speech recognitions are Google assistant, Siri, Cortana, Alexa, etc.
7. **Differentiate between supervised, unsupervised and reinforcement learning with example.**
8. **Supervised Learning:**

Supervised Learning is a machine learning method that needs supervision similar to the student-teacher relationship. In supervised Learning, a machine is trained with well-labeled data, which means some data is already tagged with correct outputs. So, whenever new data is introduced into the system, supervised learning algorithms analyze this sample data and predict correct outputs with the help of that labeled data.

It is classified into two different categories of algorithms. These are as follows:

* **Classification:** It deals when output is in the form of a category such as Yellow, blue, right or wrong, etc.
* **Regression:** It deals when output variables are real values like age, height, etc.

This technology allows us to collect or produce data output from experience. It works the same way as humans learn using some labeled data points of the training set. It helps in optimizing the performance of models using experience and solving various complex computation problems.

1. **Unsupervised Learning:**

Unlike supervised learning, unsupervised Learning does not require classified or well-labeled data to train a machine. It aims to make groups of unsorted information based on some patterns and differences even without any labelled training data. In unsupervised Learning, no supervision is provided, so no sample data is given to the machines. Hence, machines are restricted to finding hidden structures in unlabeled data by their own.

It is classified into two different categories of algorithms. These are as follows:

* **Clustering**: It deals when there is a requirement of inherent grouping in training data, e.g., grouping students by their area of interest.
* **Association:** It deals with the rules that help to identify a large portion of data, such as students who are interested in ML and also interested in AI.

1. **Reinforcement learning:**

Reinforcement learning is defined as a feedback-based machine learning method that does not require labeled data. In this learning method, an agent learns to behave in an environment by performing the actions and seeing the results of actions. Agents can provide positive feedback for each good action and negative feedback for bad actions. Since, in reinforcement learning, there is no training data, hence agents are restricted to learn with their experience only.

# **Unit 2: Supervised Learning**

1. **What is gradient descent algorithm? Explain its variants.**

Gradient descent is an optimization algorithm used to minimize some convex function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient. In machine learning, we use gradient descent to update the parameters of our model. Parameters refer to coefficients in Logistic Regression and weights in neural networks. t is basically used for updating the parameters of the learning model.

Types of Gradient Descent:

**Batch Gradient Descent:** This is a type of gradient descent which processes all the training examples for each iteration of gradient descent. But if the number of training examples is large, then batch gradient descent is computationally very expensive. Hence if the number of training examples is large, then batch gradient descent is not preferred. Instead, we prefer to use stochastic gradient descent or mini-batch gradient descent.

Algorithm:

*for i in range(#epochs):*

*grad = evaluategradient(data, para)*

*para= para – learning\_rate \* grad*

**Stochastic Gradient Descent:** This is a type of gradient descent which processes 1 training example per iteration. Hence, the parameters are being updated even after one iteration in which only a single example has been processed. Hence this is quite faster than batch gradient descent. But again, when the number of training examples is large, even then it processes only one example which can be additional overhead for the system as the number of iterations will be quite large.

Algorithm:

*for i in range(#epochs):*

*np.random.shuffle(data)*

*for d in data:*

*grad = compute\_gradient(d, params)*

*params = params — learning\_rate \* grad*

**Mini Batch gradient descent:** This is a type of gradient descent which works faster than both batch gradient descent and stochastic gradient descent. Here b examples where b<m are processed per iteration. So even if the number of training examples is large, it is processed in batches of b training examples in one go. Thus, it works for larger training examples and that too with lesser number of iterations.

Algorithm:

*for i in range(#epochs):*

*np.random.shuffle(data)  
 for batch in data:  
 grad = compute\_gradient(batch, params)  
 params = params — learning\_rate \* grad*

1. **Explain why locally weighted linear regression are used for local window of data with suitable derivations.**

Answer

1. **Derive update rule for the training of ML model using gradient descent algorithm.**

Answer

1. **How the prediction can be done using locally weighted linear regression.**

Answer

1. **Overfitting and Underfitting. How the training error can be measured and minimized?**

**Underfitting:** A statistical model or a machine learning algorithm is said to have underfitting when it cannot capture the underlying trend of the data, i.e., it only performs well on training data but performs poorly on testing data. (It’s just like trying to fit undersized pants!) Underfitting destroys the accuracy of our machine learning model. Its occurrence simply means that our model or the algorithm does not fit the data well enough. It usually happens when we have fewer data to build an accurate model and also when we try to build a linear model with fewer non-linear data. In such cases, the rules of the machine learning model are too easy and flexible to be applied on such minimal data and therefore the model will probably make a lot of wrong predictions. Underfitting can be avoided by using more data and also reducing the features by feature selection.

In a nutshell, Underfitting refers to a model that can neither performs well on the training data nor generalize to new data.

**Reasons for Underfitting:**

* High bias and low variance
* The size of the training dataset used is not enough.
* The model is too simple.
* Training data is not cleaned and also contains noise in it.

**Techniques to reduce underfitting:**

* Increase model complexity
* Increase the number of features, performing feature engineering
* Remove noise from the data.
* Increase the number of epochs or increase the duration of training to get better results.

**Overfitting:** A statistical model is said to be overfitted when the model does not make accurate predictions on testing data. When a model gets trained with so much data, it starts learning from the noise and inaccurate data entries in our data set. And when testing with test data results in High variance. Then the model does not categorize the data correctly, because of too many details and noise. The causes of overfitting are the non-parametric and non-linear methods because these types of machine learning algorithms have more freedom in building the model based on the dataset and therefore they can really build unrealistic models. A solution to avoid overfitting is using a linear algorithm if we have linear data or using the parameters like the maximal depth if we are using decision trees.

In a nutshell, Overfitting is a problem where the evaluation of machine learning algorithms on training data is different from unseen data.

**Reasons for Overfitting are as follows:**

* High variance and low bias
* The model is too complex
* The size of the training data

**Techniques to reduce overfitting:**

* Increase training data.
* Reduce model complexity.
* Early stopping during the training phase (have an eye over the loss over the training period as soon as loss begins to increase stop training).
* Ridge Regularization and Lasso Regularization
* Use dropout for neural networks to tackle overfitting.