**Gradient Descent:**

The next important concept needed to understand linear regression is gradient descent. Gradient descent is a method of updating a\_0 and a\_1 to reduce the cost function (MSE). The idea is that we start with some values for a\_0 and a\_1 and then we change these values iteratively to reduce the cost. Gradient descent helps us on how to change the Values.

Chart

Description automatically generated

Fig: Gradient Descent

You may be wondering how to use gradient descent to update a\_0 and a\_1. To update a\_0 and a\_1, we take gradients from the cost function. To find these gradients, we take partial derivatives with respect to a\_0 and a\_1. Now, to understand how the partial derivatives are found below you would require some calculus but if you don’t, it is alright.

You can take it as it is.

Text

Description automatically generated

The partial derivate are the gradients and they are used to update the values of a\_0 and a\_1. Alpha is the learning rate which is a hyper parameter that you must specify. A smaller learning rate could get you closer to the minima but takes more time to reach the minima, a larger learning rate converges sooner but there is a chance that you could overshoot the minima. The best way to understand linear regression is to relive this experience of childhood. Let us say, you ask a child in fifth grade to arrange people in his class by increasing order of weight, without asking them their weights! What do you think the child will do? She/he would likely look (visually analyze) at the height and build of people and arrange them using a combination of these visible parameters. This is linear regression in real life! The child has actually figured out that height and build would be correlated to the weight by a relationship, which looks like the equation above. In the equation, these coefficients a and b are derived based on minimizing the sum of squared difference of distance between data points and regression line.

**2) Unsupervised Learning:**

In this algorithm, we do not have any target or outcome variable to predict / estimate. It is used for clustering population in different groups, which is widely used for segmenting customers in different groups for specific intervention. Unsupervised learning algorithms are extremely powerful tools for analyzing data and for identifying patterns and trends. They are most commonly used for clustering similar input into logical groups. Unsupervised learning algorithms include Kmeans, Random Forests, and Hierarchical clustering and so on.

Examples of Unsupervised Learning: Apriori algorithm, K-means

**3. Reinforcement Learning:**

Using this algorithm, the machine is trained to make specific decisions. It works this way: the machine is exposed to an environment where it trains itself continually using trial and error. This machine learns from past experience and tries to capture the best possible knowledge to make accurate business decisions.

Example of Reinforcement Learning: Markov Decision Process.

Similarly, there are four categories of machine learning algorithms as shown below

* Supervised learning algorithm
* Unsupervised learning algorithm
* Semi-supervised learning algorithm
* Reinforcement learning algorithm

However, the most commonly used ones are supervised and unsupervised learning. Purpose of Machine Learning: Machine learning can be seen as a branch of AI or Artificial Intelligence, since, the ability to change experience into expertise or to detect patterns in complex data is a mark of human or animal intelligence. As a field of science, machine learning shares common concepts with other disciplines such as statistics, information theory, game theory, and optimization. As a subfield of information technology, its objective is to program machines so that they will learn. However, it is to be seen that, the purpose of machine learning is not building an automated duplication of intelligent behavior, but using the power of computers to complement and supplement human intelligence.