Machine Learning Diploma

Level3: Machine Learning

Session1

AMIT

Agenda

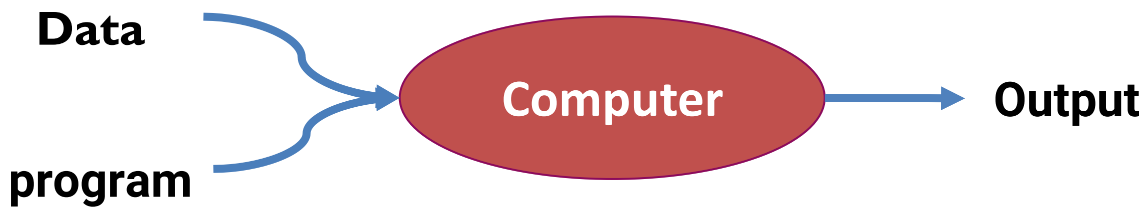
- ➔ Review for different types of Machine Learning
- ➔ Classification Vs Regression
- ➔ Linear Regression
- ➔ Gradient Descent and Cost Function
- ➔ Implementation of Linear Regression

AMIT

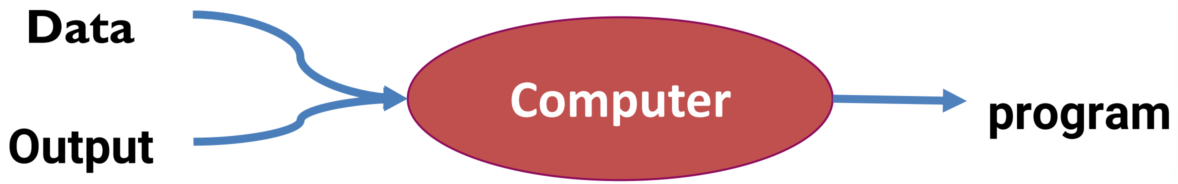
Introduction

Machine learning

- It’s an application of AI
- Computers observe and analyze
- Predict based on previous patterns

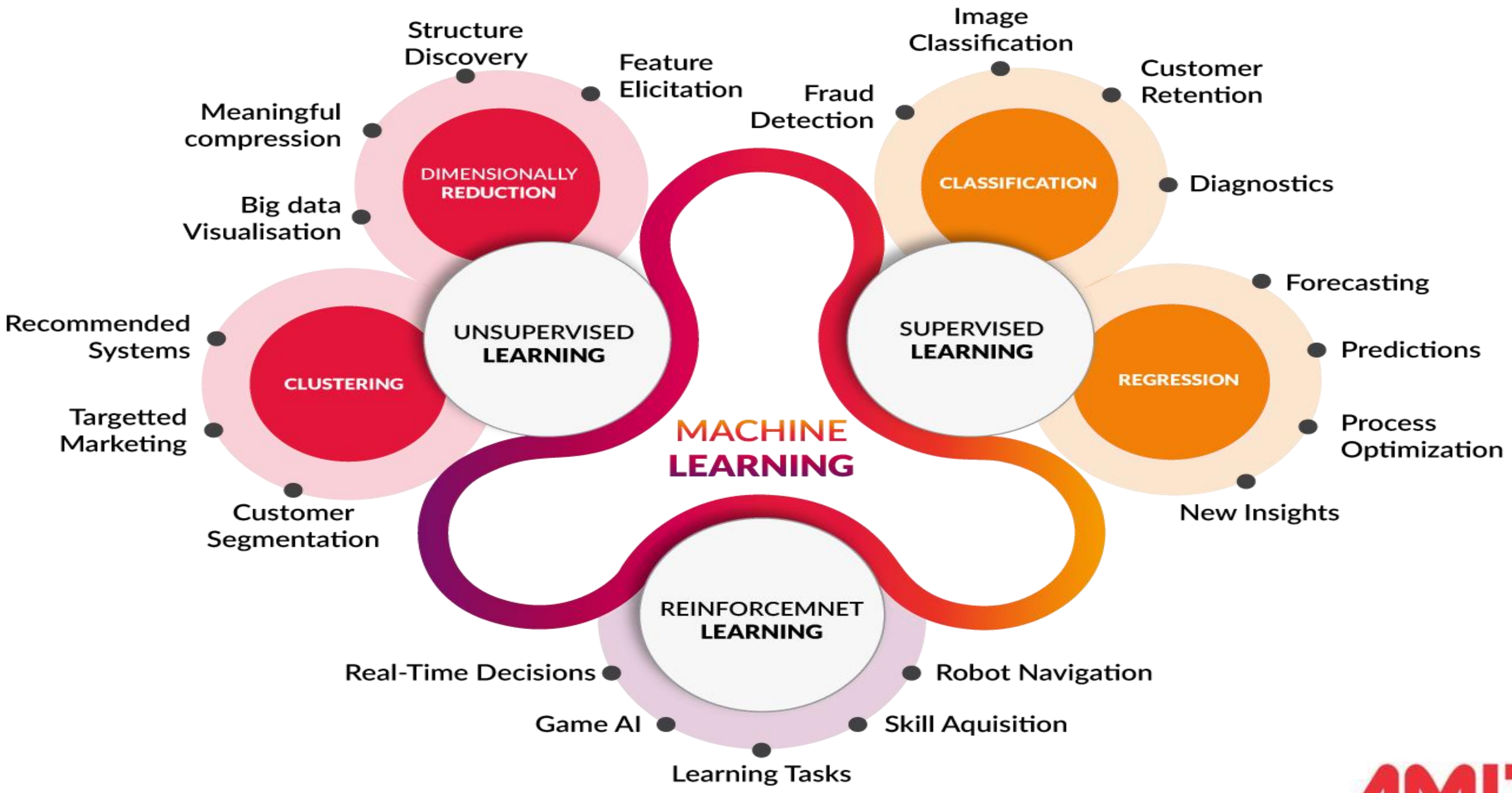


Traditional Programming chart



Machine learning chart

1. Review for different types of Machine Learning


MACHINE LEARNING
UNSUPERVISED LEARNING
SUPERVISED LEARNING
REINFORCEMNET LEARNING
DIMENSIONALLY REDUCTION
CLUSTERING
CLASSIFICATION
REGRESSION
Structure Discovery
Feature Elicitation
Meaningful compression
Big data Visualisation
Recommended Systems
Targetted Marketing
Customer Segmentation
Image Classification
Customer Retention
Fraud Detection
Diagnostics
Forecasting
Predictions
Process Optimization
New Insights
Real-Time Decisions
Robot Navigation
Game AI
Skill Aquisition
Learning Tasks

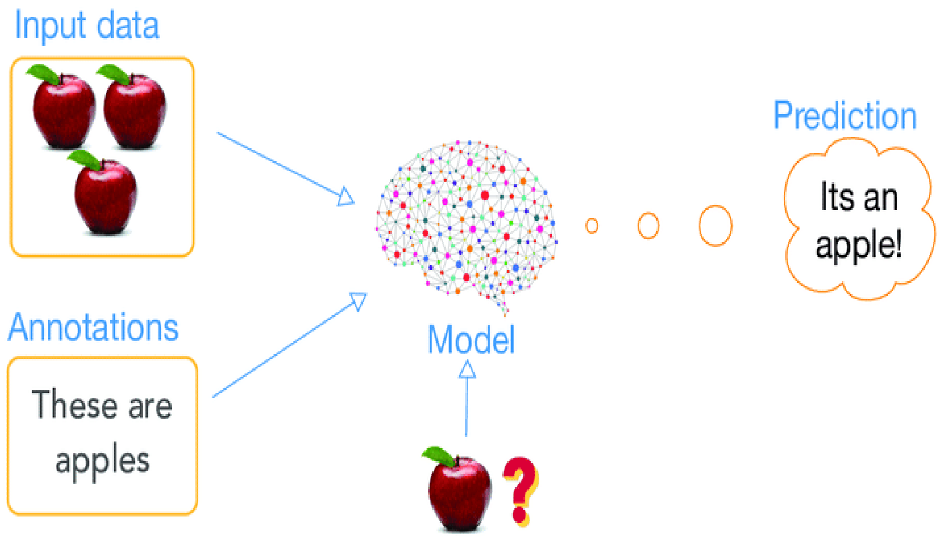
AMIT

Definition:

- ➔ Machine learning is a field of data science that focuses on designing algorithms that can learn from and make predictions on data.
- ➔ Machine Learning Types:
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning

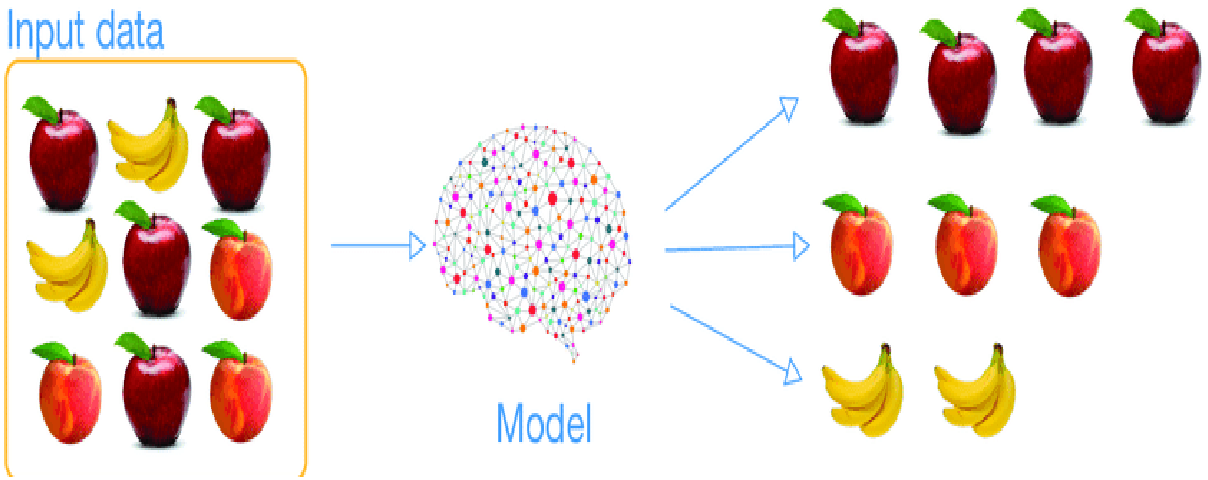
Supervised learning

Supervised learning, an AI system is presented with data which is labeled, which means that each data tagged with the correct label.



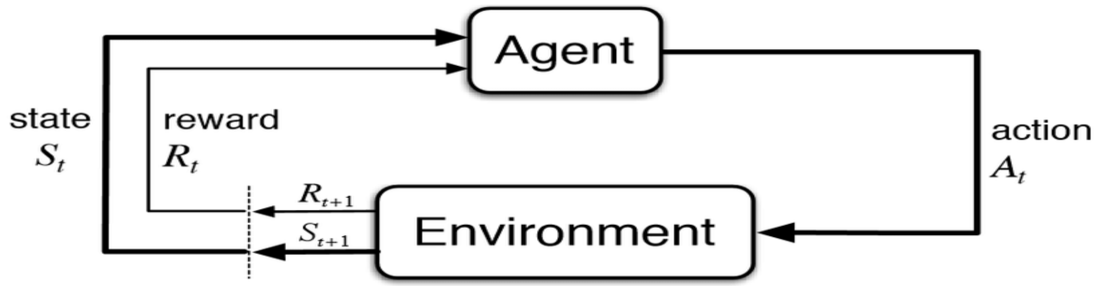
unsupervised learning

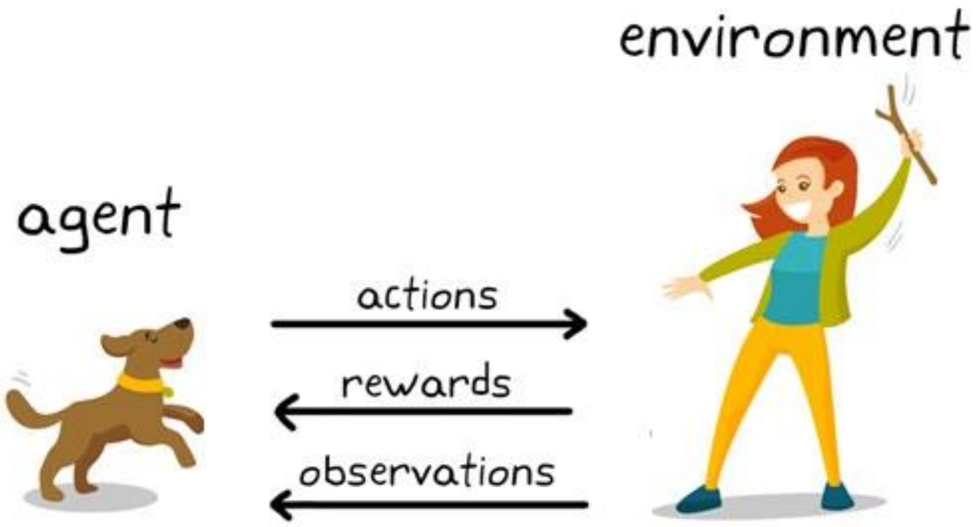
Supervised learning, AI system is presented with unlabeled, uncategorized data and the system's algorithms act on the data without prior training.



Reinforcement learning

Reinforcement learning is a type of machine learning in which a computer learns to perform a task through repeated trial-and-error interactions with a dynamic environment.





AMIT

2. Classification Vs Regression

Regression:

- ➔ Regression is a technique to model output value with the help of independent predictors.
- ➔ Regression is used to predict a relationship between a dependent variable and independent variables.

Regression Problem try to learn the real numerical value of the class, such as “dollars” or “weight” or “size”.

AMIT

Classification:

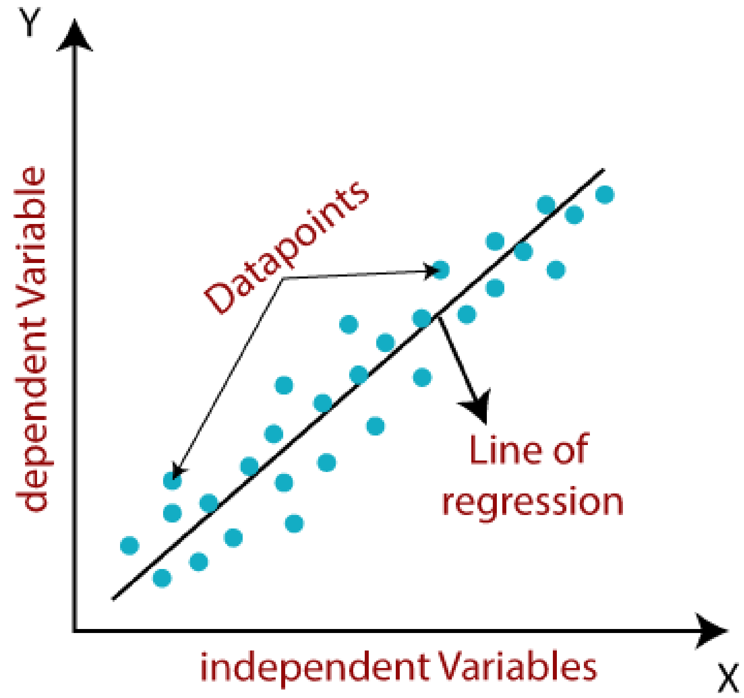
- ➔ Classification is a supervised learning concept which basically categorizes a set of data into classes.
- ➔ It is a predictive modelling problem where a class label is predicted for a given example of input data.

Classification problem try to learn categorical class.
such as “red” or “blue” or “yellow”

3. Linear Regression

Linear Regression:

- ➔ Linear Regression describes a linear relationship between an independent variable and the dependent variable.
- ➔ The main idea is to predict the best fit line with the help of given data points.
- ➔ The best fit line will have minimum error depicted by the distance between the points and the line.



$$Y = a_0 + a_1 X$$

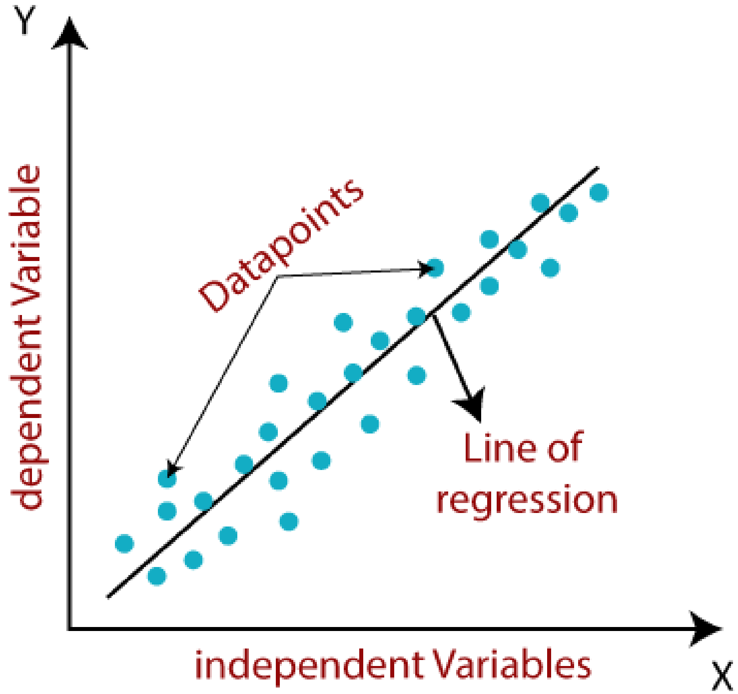
4. Gradient Descent and Cost Function

Cost Function:

- Cost Function helps us to get the best values for a_0 and a_1, in order to get the best fit line for data points.
- The difference between the predicted values and the actual values gives the error.
- We want to minimize this error "mean squared error function"

$$\text{Minimize } \frac{1}{n}\sum_{i=1}^{n}(predict_i - y_i)^2$$

$$J = \frac{1}{n}\sum_{i=1}^{n}(predict_i - y_i)^2$$



AMIT

Cost Function (evaluation metrics) for Regression example:

➔ Mean squared Error **(MSE)**

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

➔ Mean absolute Error **(MAE)**

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \hat{y}_j\right|$$

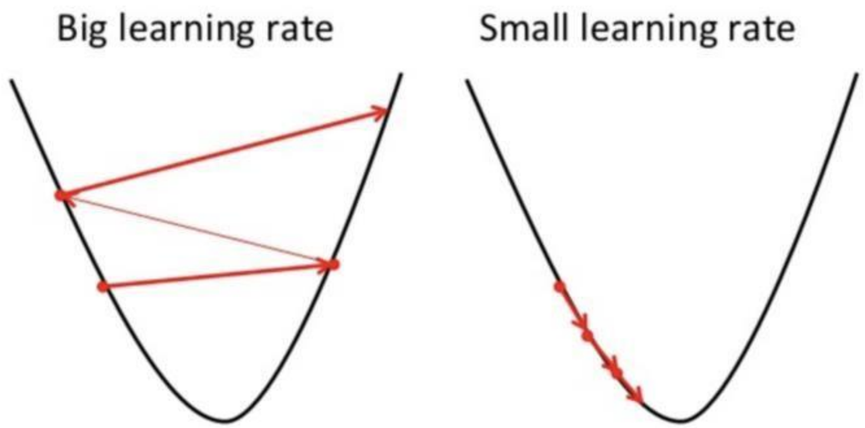
Gradient Descent:

- ➔ Gradient Descent helps us to reduce the cost function to reach minima.
- ➔ Start with some values of a_0 and a_1.
- ➔ Change them iteratively until reaching a minimum for cost function **J**.

- ➔ Learning Rate **α** determines how fast we are going to reach the minima.

$$a_0 = a_0 - \alpha \,.\, \frac{2}{n}\sum_{i=1}^{n}(predict_i - y_i)$$

$$a_1 = a_1 - \alpha \,.\, \frac{2}{n}\sum_{i=1}^{n}(predict_i - y_i).\, x_i$$

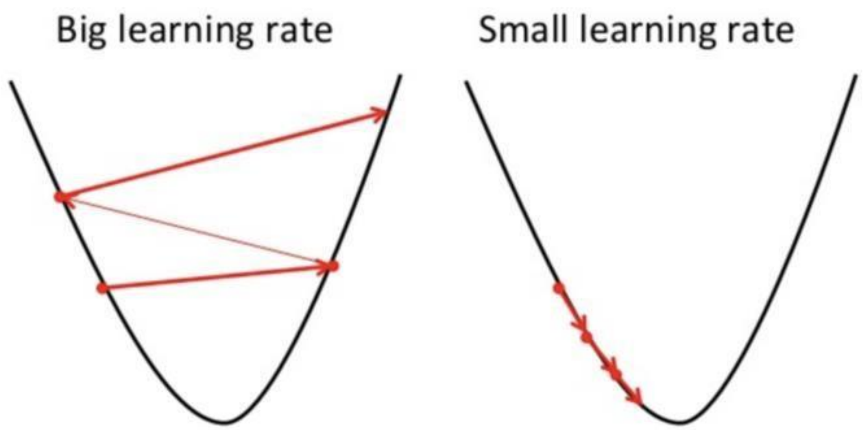


Gradient Descent:

- ➔ To calculate these steps or gradients from cost function, we take gradients (partial derivatives) of **a_0** and **a_1**. A little bit of calculus will do the trick here.
- ➔ Higher learning rate means less time but with a chance of missing the minima and vice versa for the lower learning rate.

$$a_0 = a_0 - \alpha \,.\, \frac{2}{n}\sum_{i=1}^{n}(predict_i - y_i)$$

$$a_1 = a_1 - \alpha \,.\, \frac{2}{n}\sum_{i=1}^{n}(predict_i - y_i).\, x_i$$



Gradient Descent find new a_0 and a_1 :

$$y_{predict} = a_0 + a_1 \,.\, x$$

$$j = \frac{1}{n}\sum_{i}^{n}(predict_i \; - \; y_i)^2 \qquad = \qquad j = \frac{1}{n}\sum_{i}^{n}((a_0 \; + a_1 . x) - y_i)^2$$

Apply differentiation on *j* twice for a_o and a_1 to get D_{a_o} and D_{a_1}

$$D_{a_o} = \frac{2}{n}\sum_{i=1}^{n}(predict_i - y_i)$$

$$D_{a_1} = \frac{2}{n}\sum_{i=1}^{n}(predict_i - y_i).\; x$$

$$a_o = a_o - a \,.\, D_{ao}$$

new a_0 *old* a_0 *Learning rate*

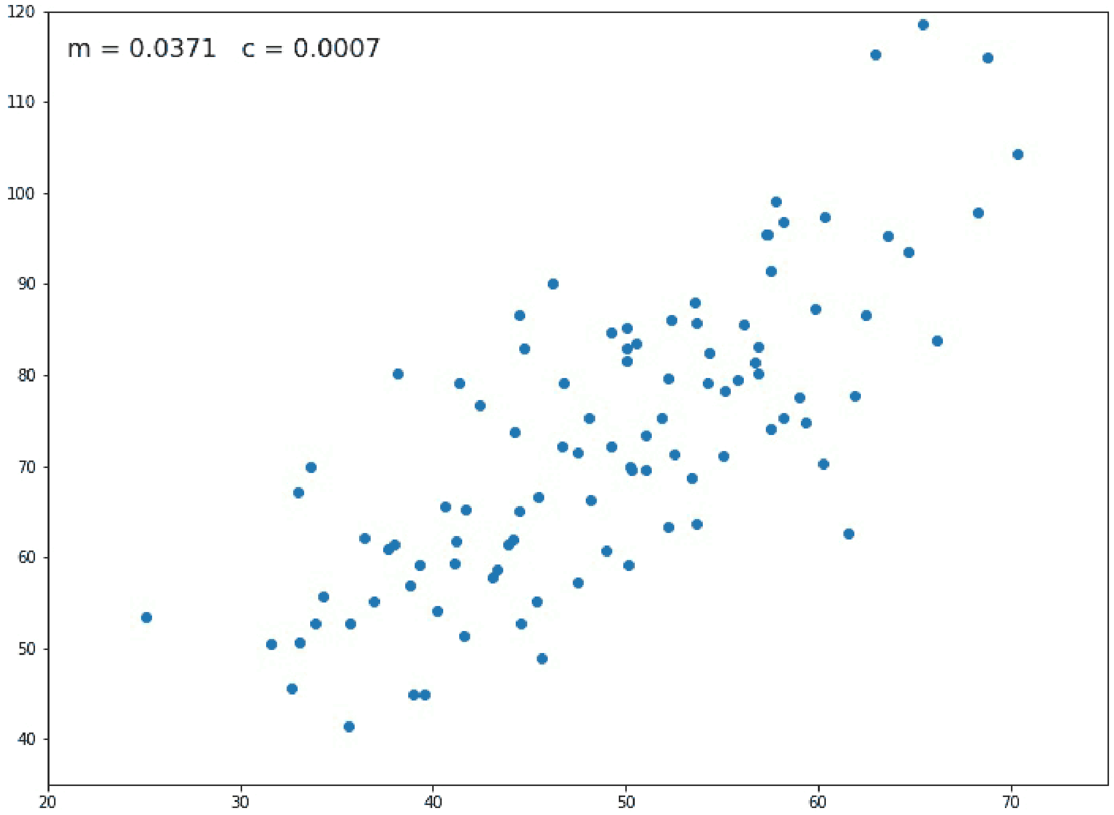
$$a_1 = a_1 - a \,.\, D_{a1}$$

new a_1 *old* a_1 *Learning rate*

repeat this for number of epochs

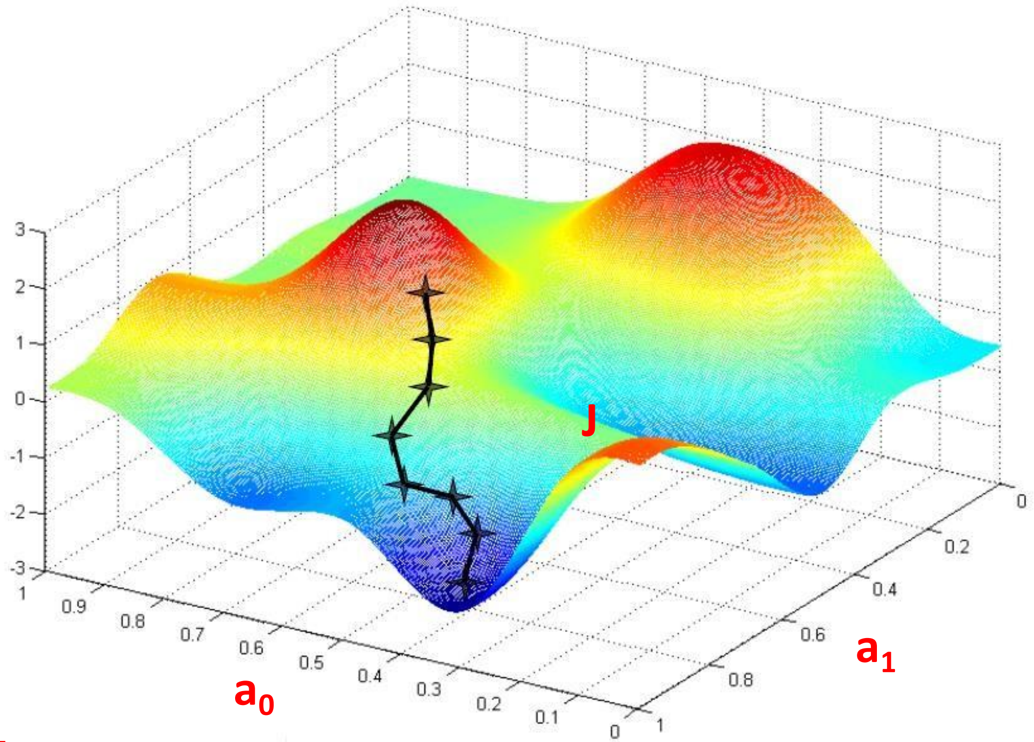
AMIT

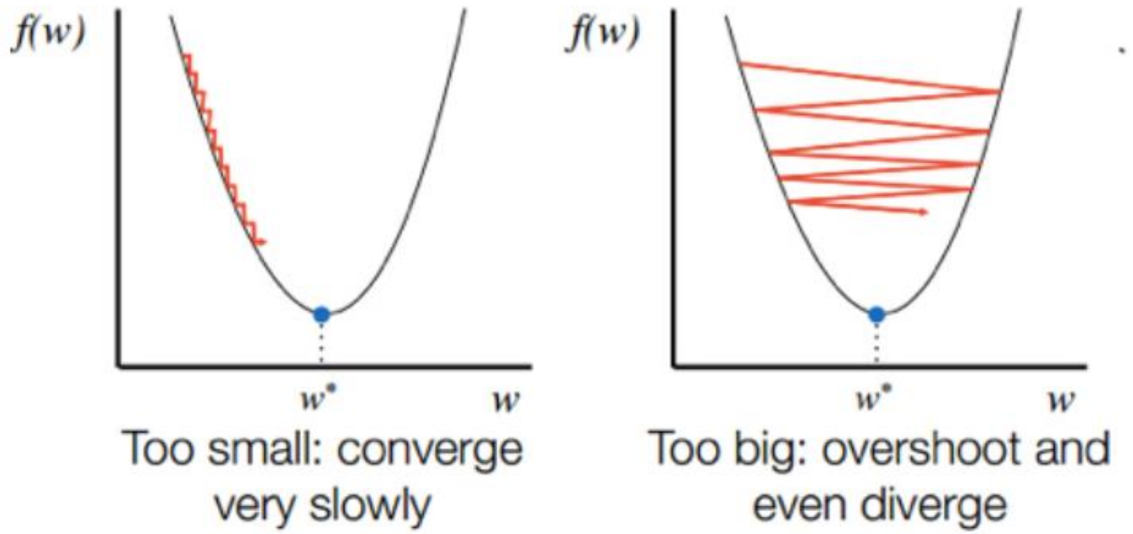
Gradient Descent:



Gradient Descent:

- ➔ If **α** is too small, gradient descent can be slow.
- ➔ If **α** is too large, gradient descent can overshoot the minima. It may fail to converge, or even diverge.





AMIT

THANK YOU!

AMIT