DATA SCIENCE INTERNSHIP

Session 8

HDLC TECHNOLOGIES

Machine Learning.

- Human's knowledge is only obtained by the experience throughout their life. For machines those knowledge is need to be fed, by collecting enormous amount of data on a certain application and fed to it, machines also obtains the knowledge in short period of time.

- There are three types of Machine Learning

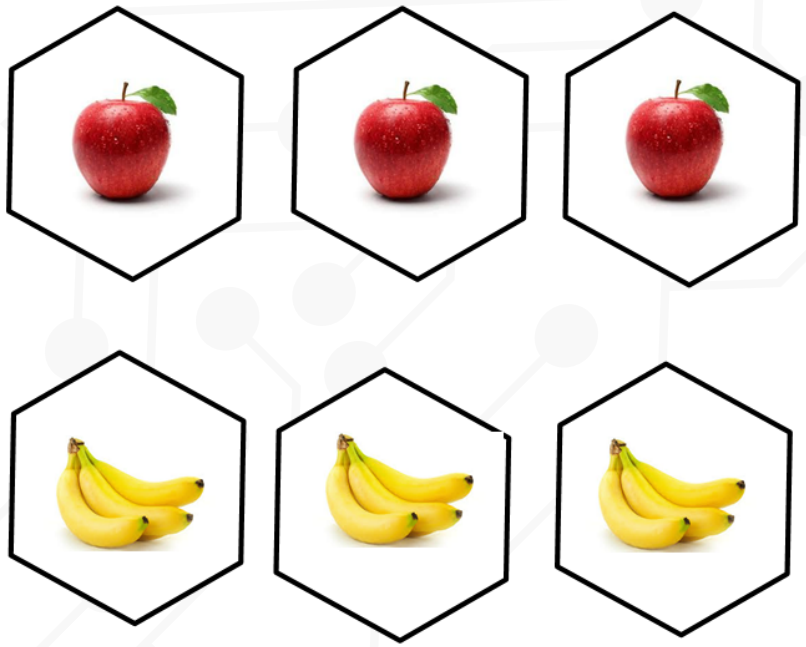
- Supervised
- Unsupervised
- Reinforcement

Supervised Learning.

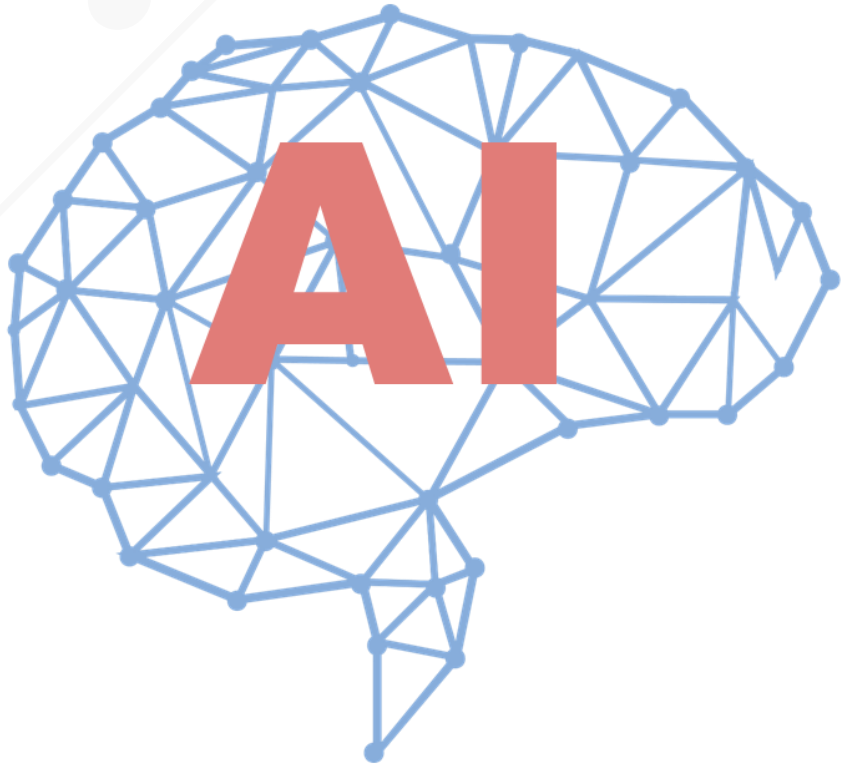
Labelled

Input Data

These are Apple

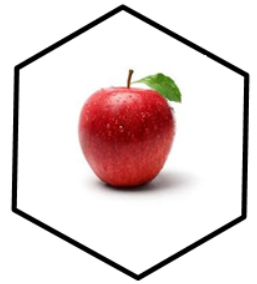


These are Banana



Model

Prediction

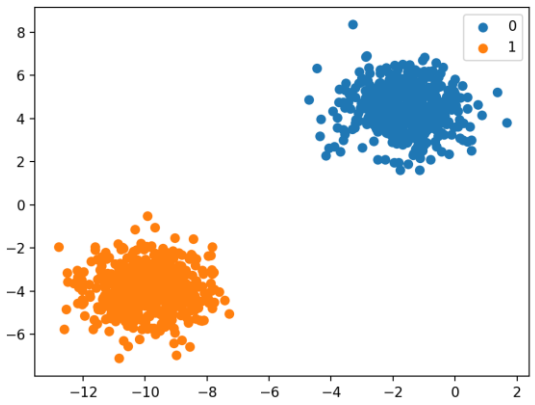


Its an Apple

SUPERVISED LEARNING.

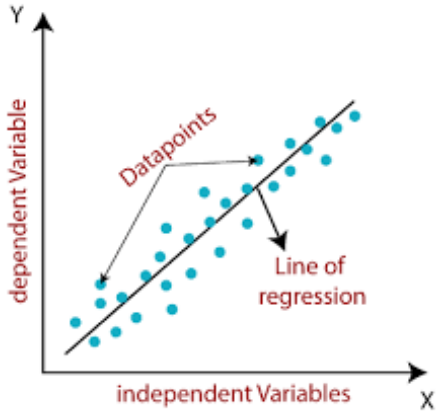
CLASSIFICATION

- Predictive modeling problem where a class label is predicted for a given example of input data.
- Given an example, classify if it is spam or not. Given a handwritten character, classify it as one of the know



PREDICTION/REGRESSION

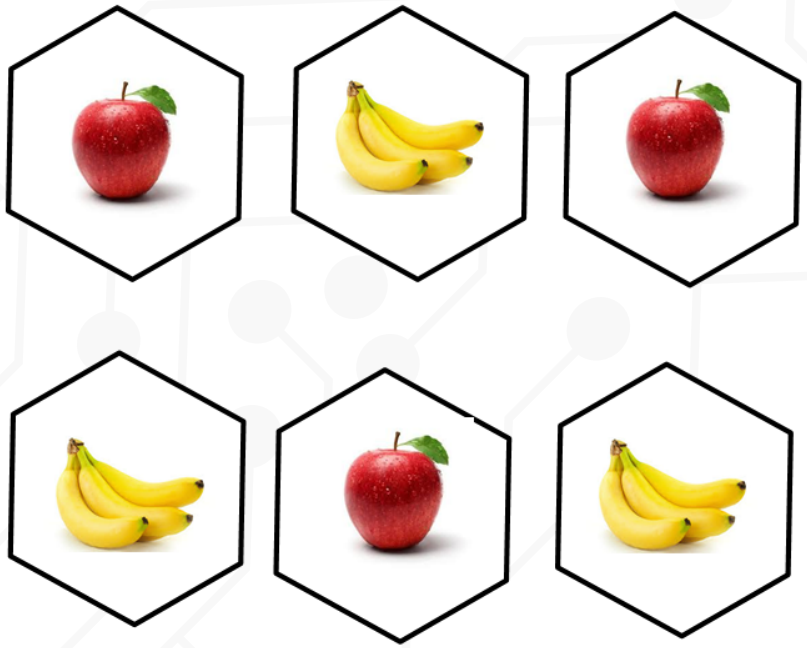
- It allow us to predict a continuous outcome variable (y) based on the value of one or multiple predictor variables (x). Briefly, the goal of regression model is to build a mathematical equation that defines y as a fu riables.
- Continuo iscrete

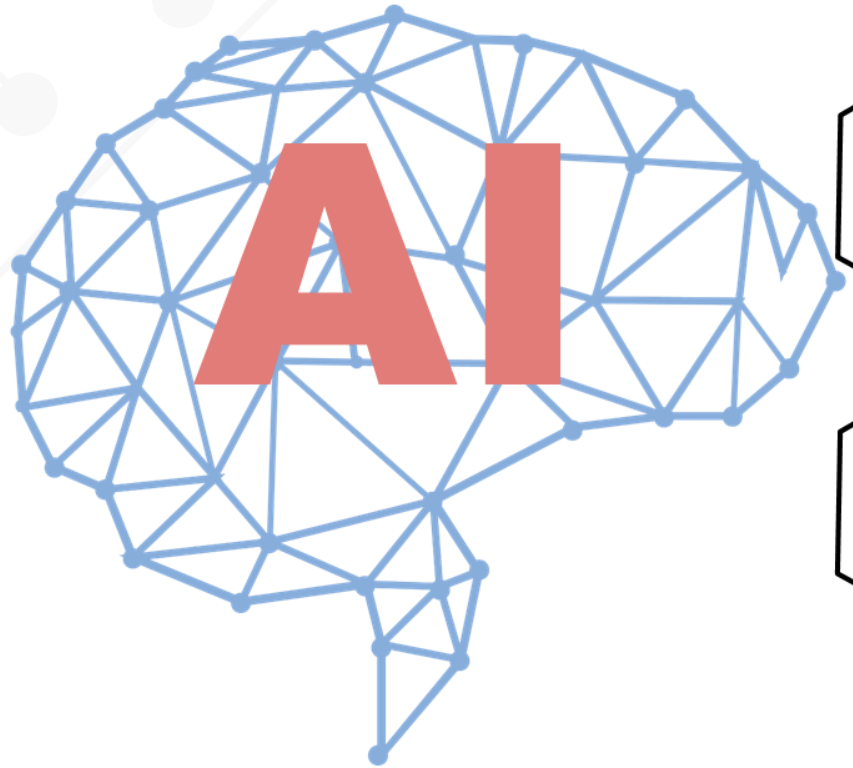


Unsupervised Learning.

Unlabelled

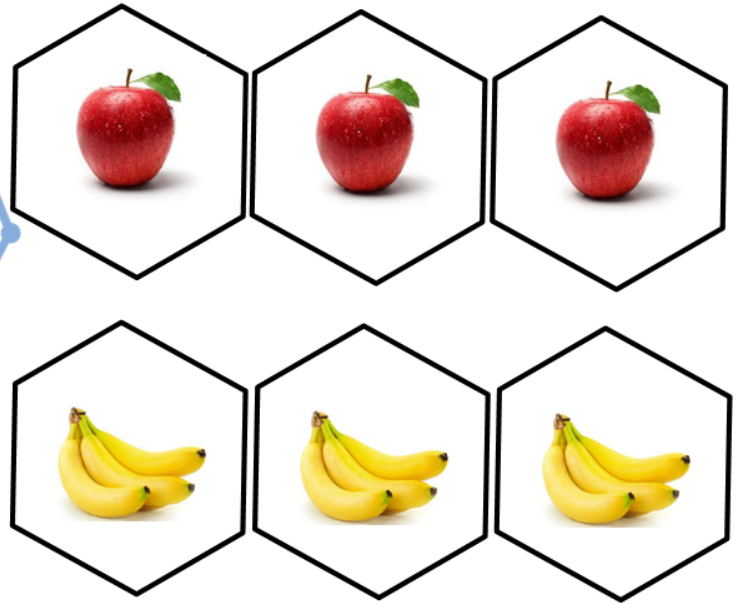
Input Data



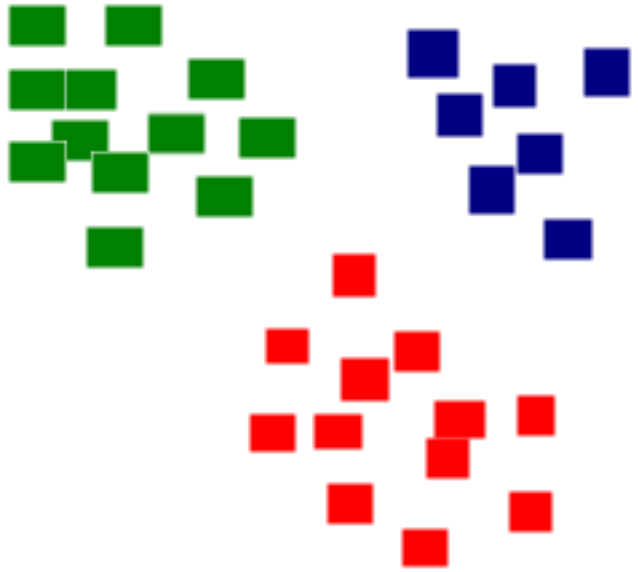


Model

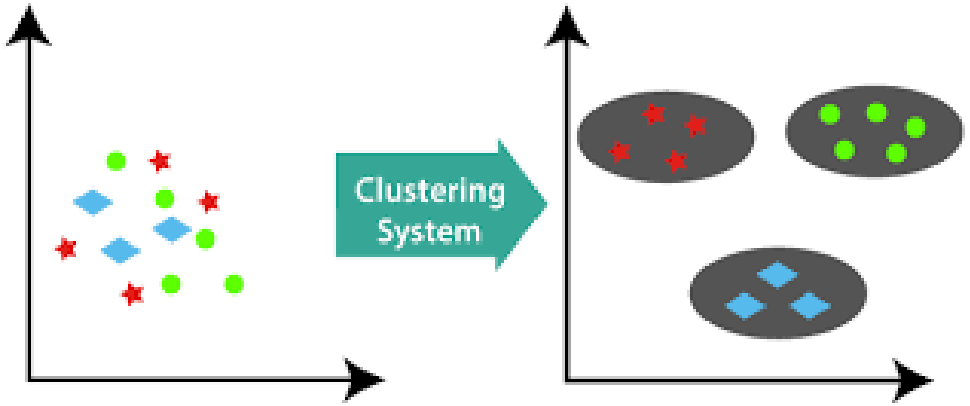
Prediction



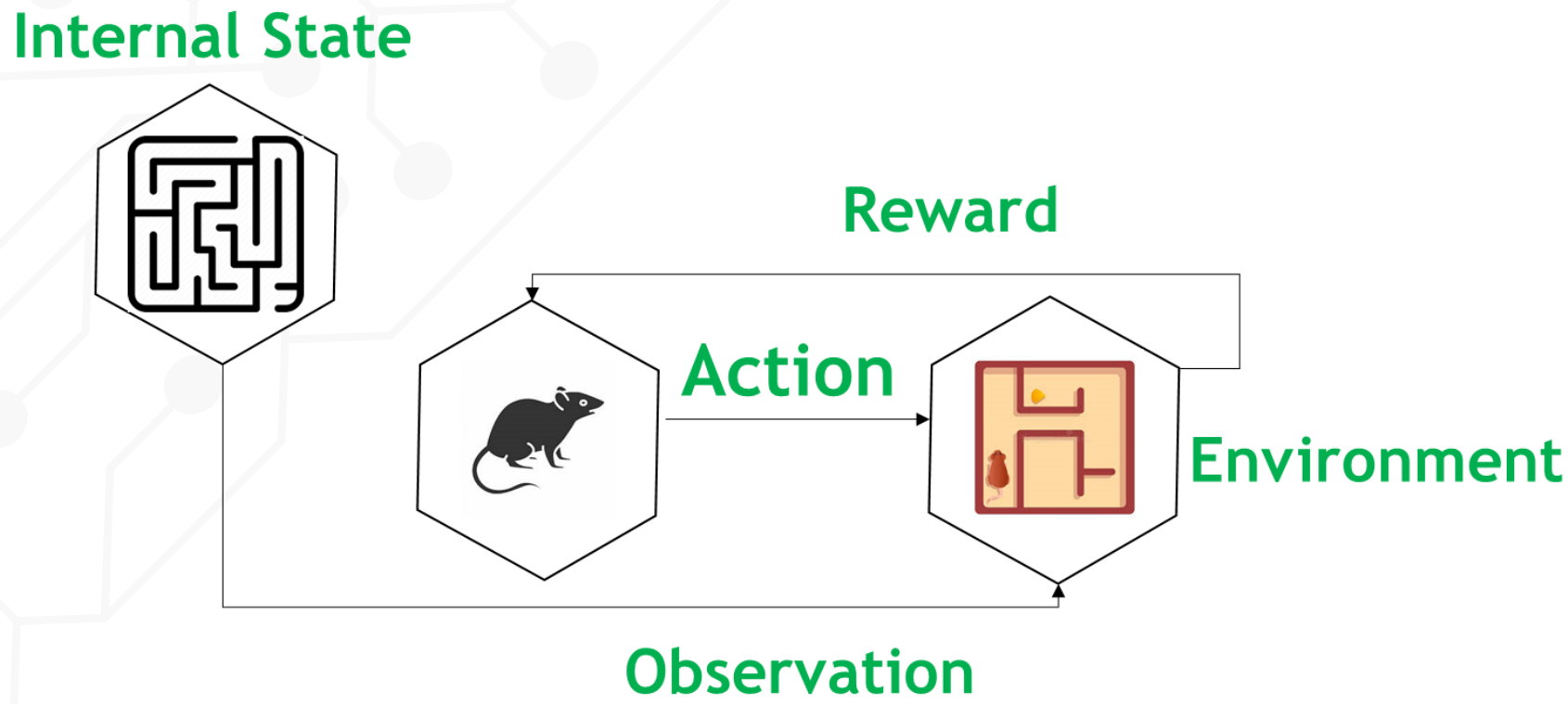
CLUSTERING



- Set of inputs is to be divided into groups.
- The groups are not known beforehand, whereas classification knows



Reinforcement Learning.



REINFORCEMENT LEARNING.

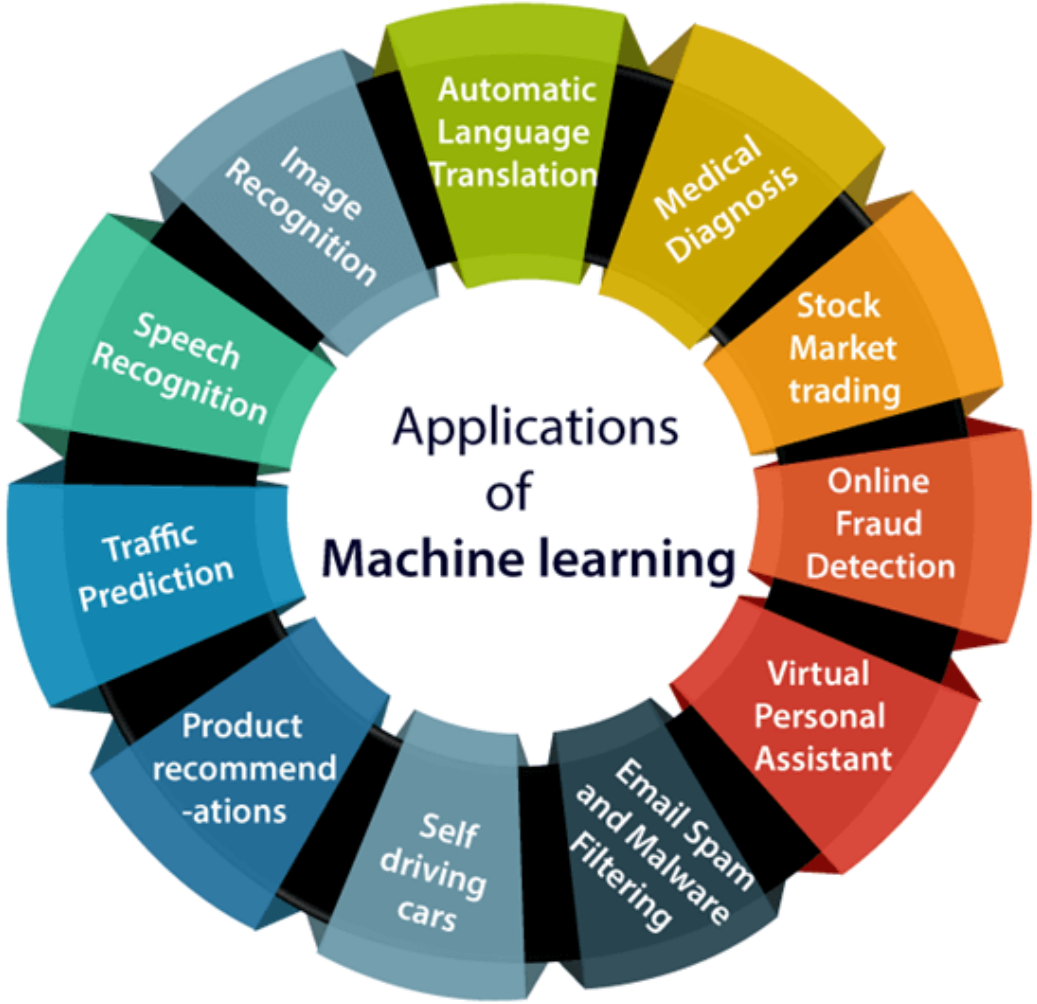
- It is about taking suitable action to maximize reward in a particular situation.
- It is employed by various software and machines to find the best possible behavior or path it should take in a specific situation.
- Output depends on the state of the current input and the next input depends on the output of the previous input

✓ Positive
✓ Negative

Machine Learning Process

- Gathering data
- Preparing that data
- Choosing a model
- Training
- Evaluation
- Hyperparameter tuning
- Prediction

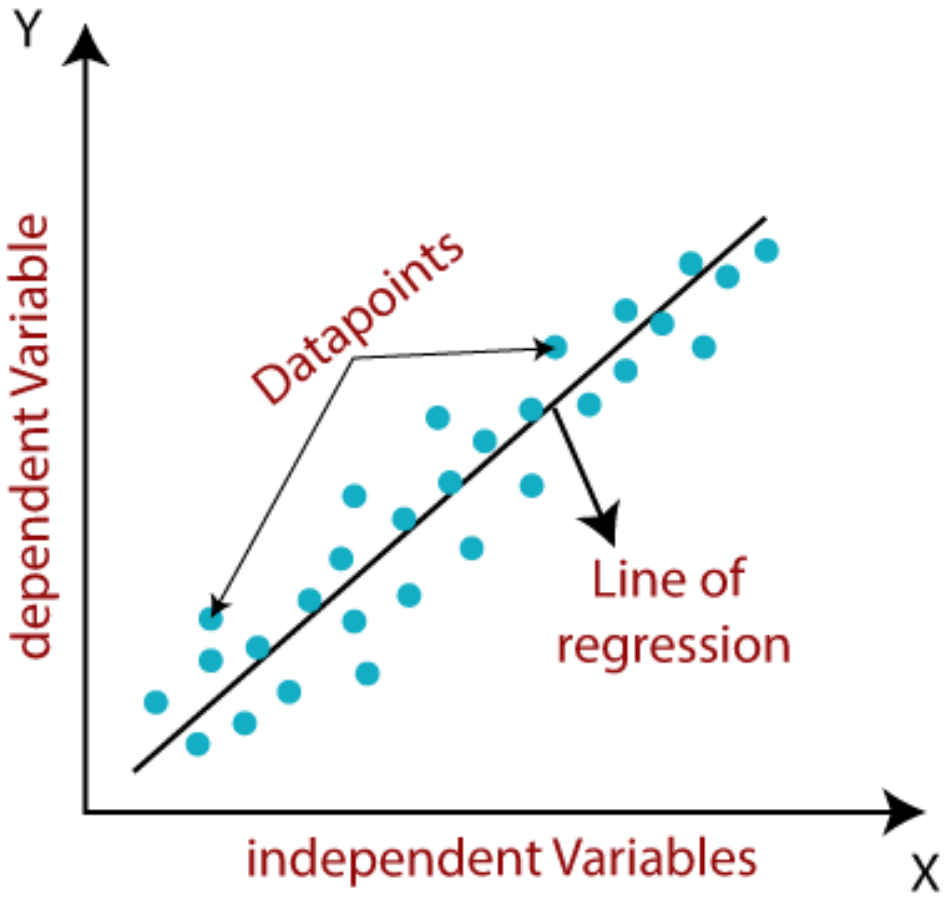
Applications of ML



Linear Regression in Machine Learning

- Simple Linear Regression is a type of Regression algorithms that models the relationship between a dependent variable and independent variables.

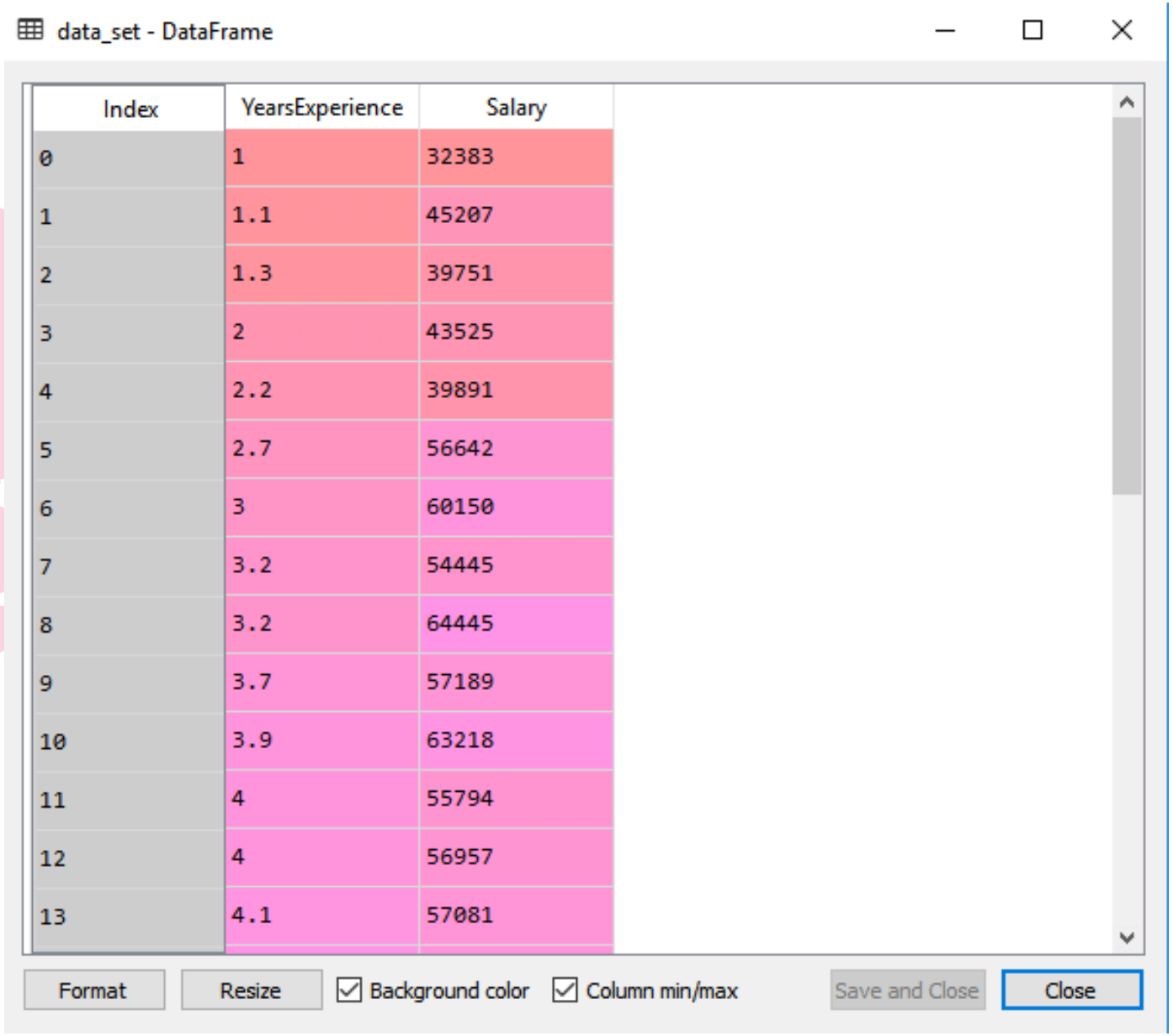
- $Y= mx+c$
- $Y=ax_1+bx_2+cx_3+d$



Steps

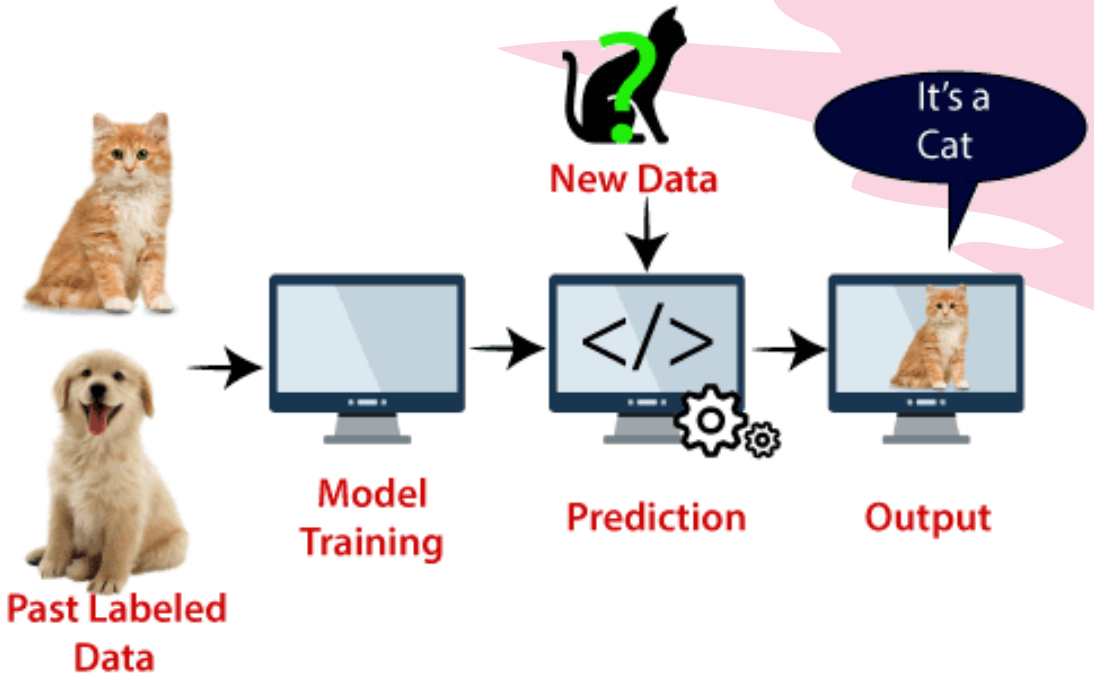
- Data Pre-processing
- Splitting the dataset into trai ning and test set.
- Fitting the Simple Linear Regression to the Training Set
- Prediction of test set result:

Salary Prediction



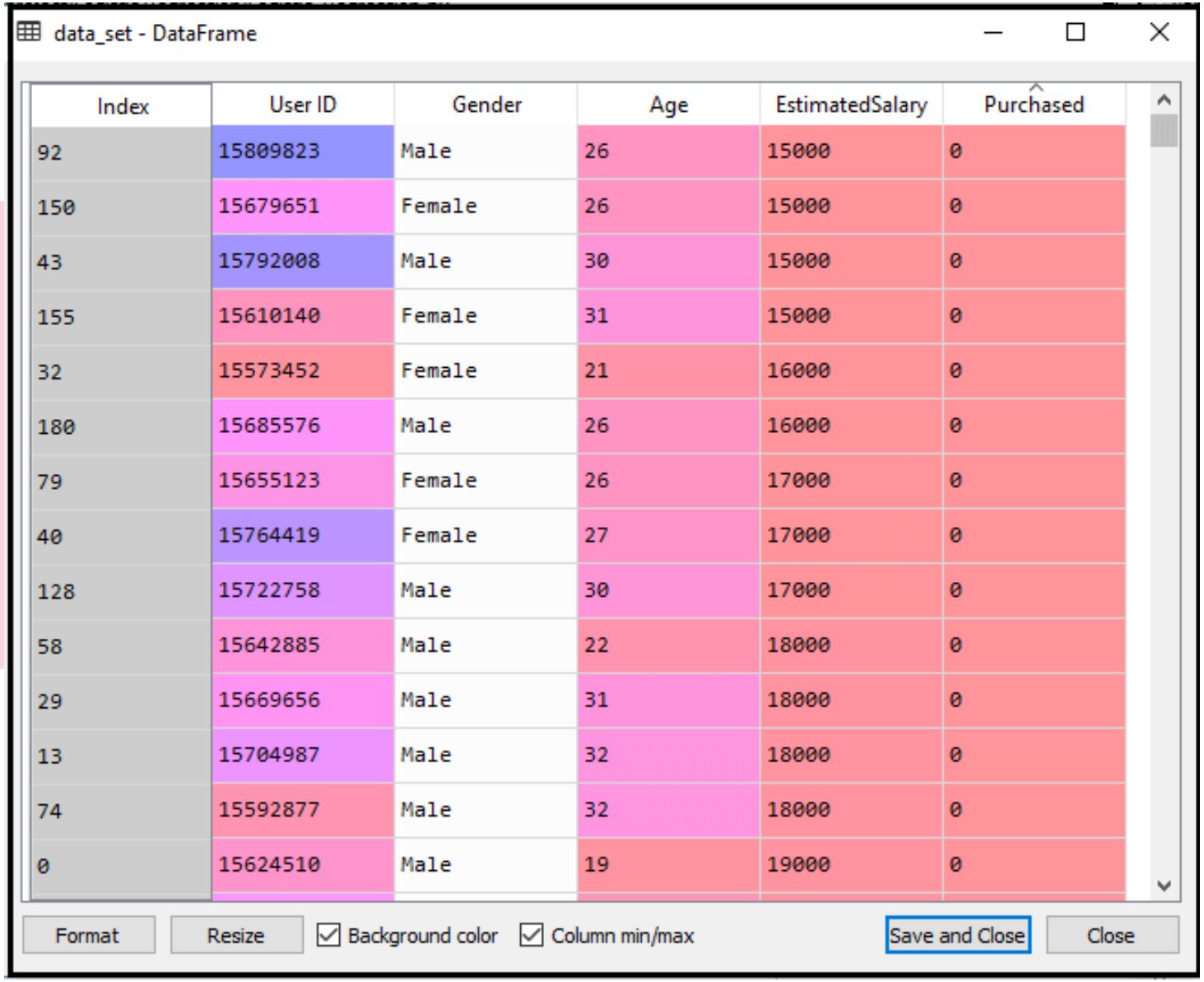
Index	YearsExperience	Salary
0	1	32383
1	1.1	45207
2	1.3	39751
3	2	43525
4	2.2	39891
5	2.7	56642
6	3	60150
7	3.2	54445
8	3.2	64445
9	3.7	57189
10	3.9	63218
11	4	55794
12	4	56957
13	4.1	57081

Support Vector Machine Algorithm



- Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

Customer purchase



Index	User ID	Gender	Age	EstimatedSalary	Purchased
92	15809823	Male	26	15000	0
150	15679651	Female	26	15000	0
43	15792008	Male	30	15000	0
155	15610140	Female	31	15000	0
32	15573452	Female	21	16000	0
180	15685576	Male	26	16000	0
79	15655123	Female	26	17000	0
40	15764419	Female	27	17000	0
128	15722758	Male	30	17000	0
58	15642885	Male	22	18000	0
29	15669656	Male	31	18000	0
13	15704987	Male	32	18000	0
74	15592877	Male	32	18000	0
0	15624510	Male	19	19000	0