KNN - K Nearest Neighbour

- KNN is a classification Algorithm
- Classification is a Supervised Learning Method

- Distance between two points Formula = $E(x,y) = \sqrt{\sum_{i=0}^{n}(x_i - y_i)^2}$

Q. We have to predict the class for the person having Age 48 and Loan Amount is 142000.

So by calculating the distance we have the nearest 5 points as mentioned

$X_1 = 48$ $Y_1 = 142000$

AGE (X_2)	Loan Amount (Y_2)	Default Status	Distance from given Point (E)	Nearest Points to given point
25	40000	N	102000.0011	
38	60000	N	82000.00002	
45	80000	Y	62000.0002	4
20	20000	N	122000.0016	
35	120000	N	22000.00057	1
52	18000	N	124000.0006	
23	90000	Y	52000.00278	3
40	62000	Y	80000	
60	100000	Y	42000.00476	2
48	220000	Y	78000.00041	5
33	10000	Y	132000.0002	

- So when we have 3-Nearest Neighbour we have [N,Y,N] as three nearest Default Status
  So based on Maximum Votes we can tell that the New person with Age 42 and Loan Amount 142000 <u>will not be a defaulter</u>.
- If we take 5-Nearest Neighbour we have [N,Y,N,Y,Y] as five nearest Default Status
  So based on Maximum Votes we can tell that the New person with Age 42 and Loan Amount 142000 <u>will be a defaulter</u>.
- We always choose Odd number of Nearest Values since if we choose even number we may come in a scenario of equal number of cases on two sides in Binary Classification.

Written by: Sharanu Ullegaddi