

SUPERVISED AND UNSUPERVISED

PRESENTED BY:
NITISH VIG

TRAIN TEST SPLIT

We split the dataset into training and testing part

Syntax : *from sklearn.model_selection import train_test_split*

xtrain, xtest, ytrain, ytest = train_test_split(x, y, test_size = 0.2, random_state = 100)

Basics

Test size = 0.2

Population = 16

Test set = 4

Train set = 12

Age	Behaviuor	Gender	Pocket Money	Hobby	Will Invest
19	Graduation	Girl	100	Cooking	Yes
20	Graduation	Boy	200	Cooking	Yes
20	Graduation	Boy	100	Jogging	No
16	Graduation	Girl	100	Jogging	No
16	Graduation	Girl	300	Jogging	No
19	Graduation	Boy	100	Jogging	Yes
20	Graduation	Girl	200	Dancing	No
19	Graduation	Girl	500	Cooking	Yes
22	Post Graduation	Girl	200	Jogging	No
22	Post Graduation	Boy	500	Dancing	Yes
22	Post Graduation	Girl	500	Dancing	Yes
19	Graduation	Boy	500	Jogging	No
19	Graduation	Girl	200	Dancing	No
16	Graduation	Girl	100	Cooking	No
22	Post Graduation	Boy	200	Dancing	No
16	Graduation	Boy	300	Dancing	Yes

SUPERVISED LEARNING

LINEAR REGRESSION

Basics

The response is continuous in nature.

Syntax : *from sklearn.linear_model import LinearRegression*

Basics

Predictors			Response
Distance (km)	Time (min.)	Car Type	Fare
10	60	0	180
6	50	1	250
6	40	0	120
4	20	1	100
8	50	0	150
3	20	1	100
3	20	0	60
0	0	0	50
8	40	1	200
16	80	0	240

LOGISTIC REGRESSION

Basics

The response is continuous in nature.

Syntax : `from sklearn.linear_model import LogisticRegression`

Basics

Predictors			Response
Distance (km)	Time (min.)	Car Type	Feedback
10	60	0	1
6	50	1	0
6	40	0	0
4	20	1	1
8	50	0	0
3	20	1	1
3	20	0	1
0	0	0	1
8	40	1	0
16	80	0	0

KMEANS

Entire dataset is sent to the training part and score is analyzed to watch over.

Syntax : *from sklearn.cluster import KMeans*

Basics

Predictors			
X1	X2	X3	X4
10	60	0	1
6	50	1	0
6	40	0	0
4	20	1	1
8	50	0	0
3	20	1	1
3	20	0	1
0	0	0	1
8	40	1	0
16	80	0	0

Basics

Predictors				Clusters
X1	X2	X3	X4	labels_
10	60	0	1	5
6	50	1	0	4
6	40	0	0	1
4	20	1	1	4
8	50	0	0	5
3	20	1	1	3
3	20	0	1	0
0	0	0	1	0
8	40	1	0	2
16	80	0	0	0