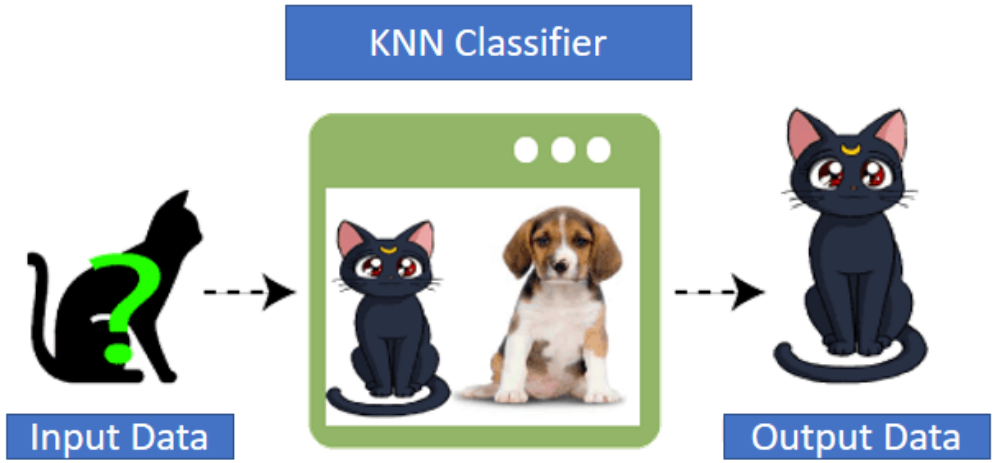
What is K-NN Algorithm?



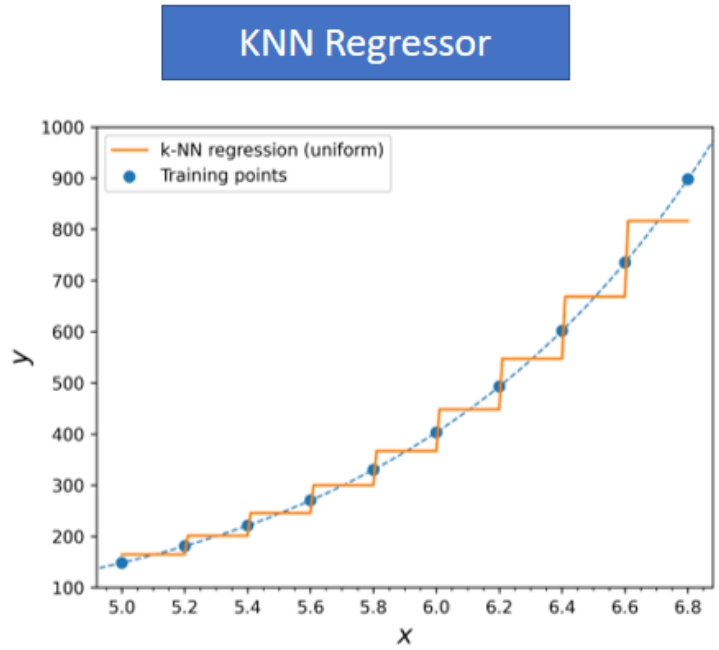


Image: javatpoint.com

Why do we need a K-NN Algorithm?

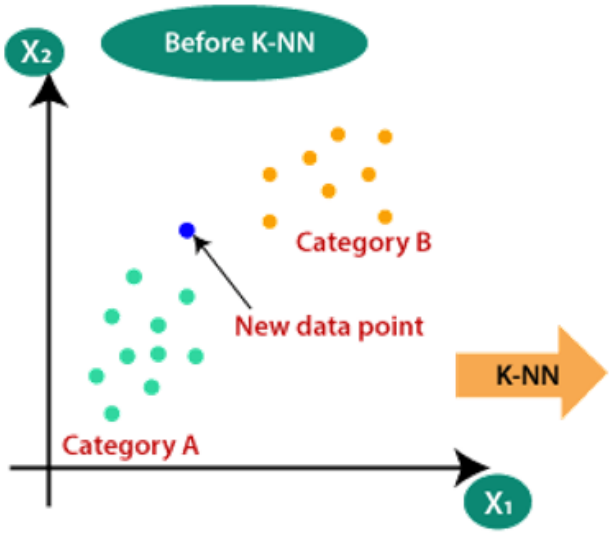


Image: javatpoint.com

Why do we need a K-NN Algorithm?

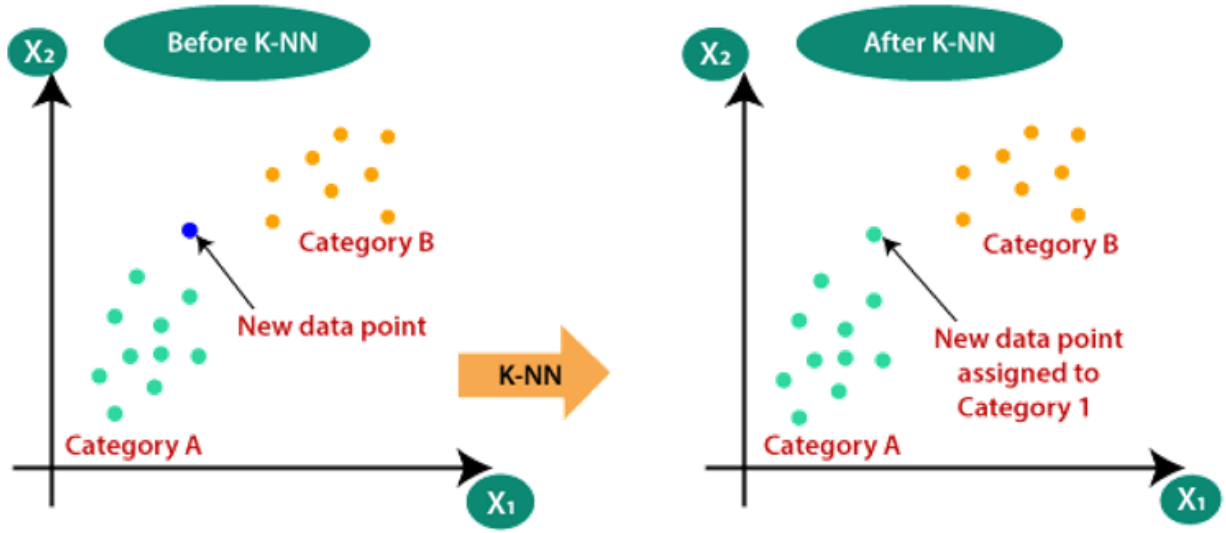


Image: javatpoint.com

Steps: K-NN Algorithm

Steps: The K-NN working can be explained on the basis of the below algorithm:

Step-0: Calculate the Euclidean distance from unknown data point.
Step-1: Sort all data point based on distance.
Step-2: Take the K nearest neighbors as per the calculated Euclidean distance.
Step-3: Among these k neighbors, count the number of the data points in each category.
Step-4: Assign the new data points to that category for which the number of the neighbor is maximum.
Step-5: Our model is ready. Let's Predict -

K-NN Algorithm: Classification

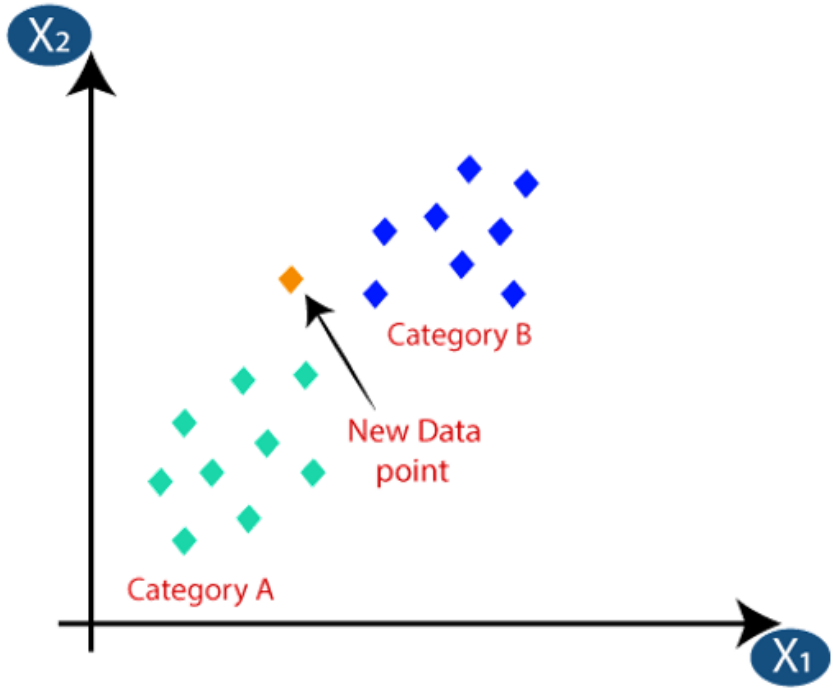
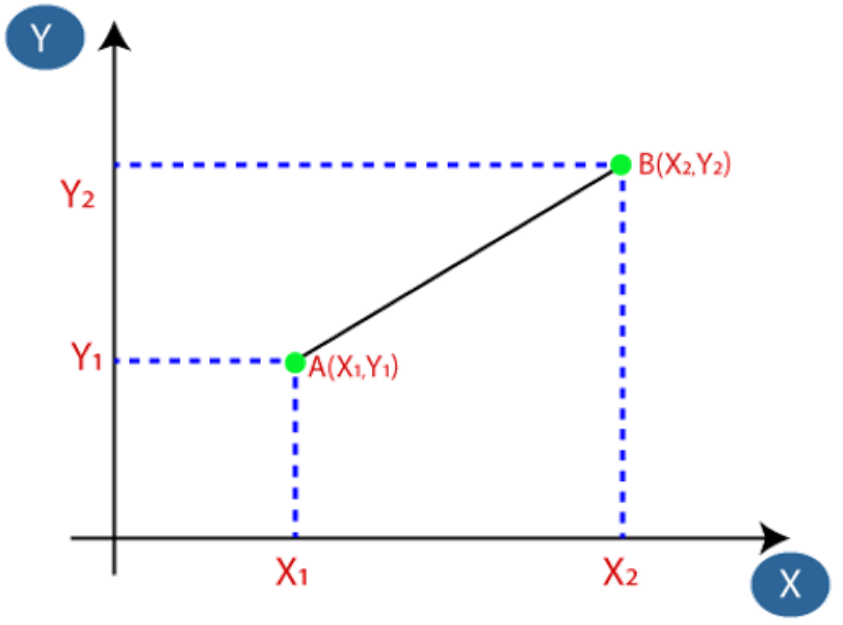


Image: javatpoint.com

K-NN Algorithm: Classification



Euclidean Distance between A_1 and B_2 = $\sqrt{(X_2-X_1)^2+(Y_2-Y_1)^2}$

Image: javatpoint.com

K-NN Algorithm: Classification

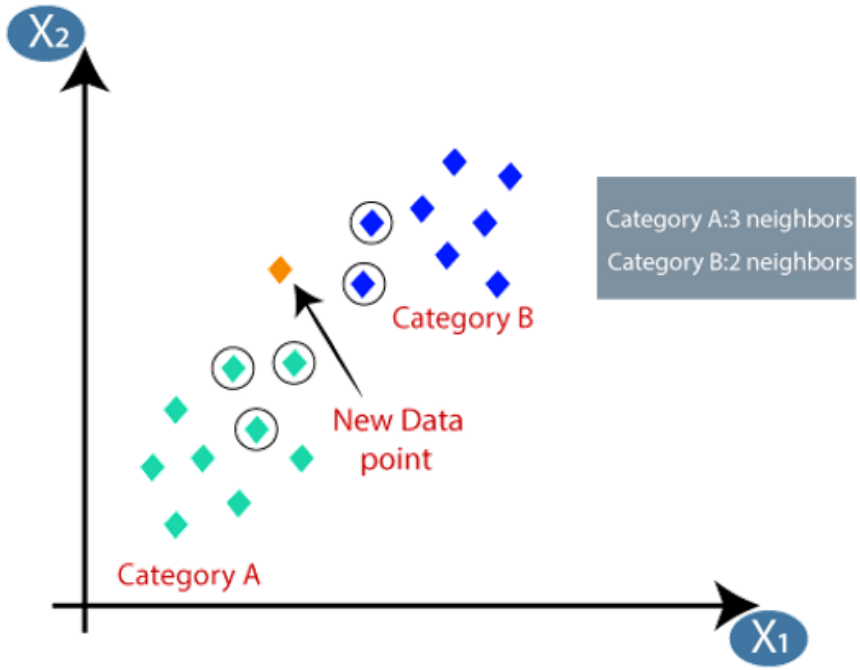
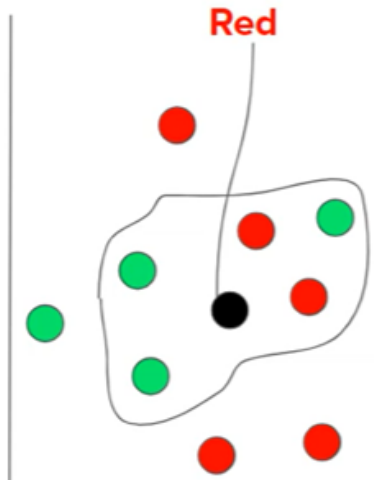


Image: javatpoint.com

K-NN Algorithm: Weights{'uniform', 'distance'}



Point	Label	Distance	Weight
(x1,y1)	Red	0.2	5
(x2,y2)	Red	0.5	2
(x3,y3)	Green	0.7	1.4
(x4,y4)	Green	1.2	0.8
(x5,y5)	Green	1.5	0.6

Calculate Weight

Based on a **Weighing Function**

Distance Increases, Weight decreases

Simplest Weighing function

1.4 + 0.8 + 0.6 = 2.8

5 + 2 = 7

K=5

$w_i = 1/d_i$

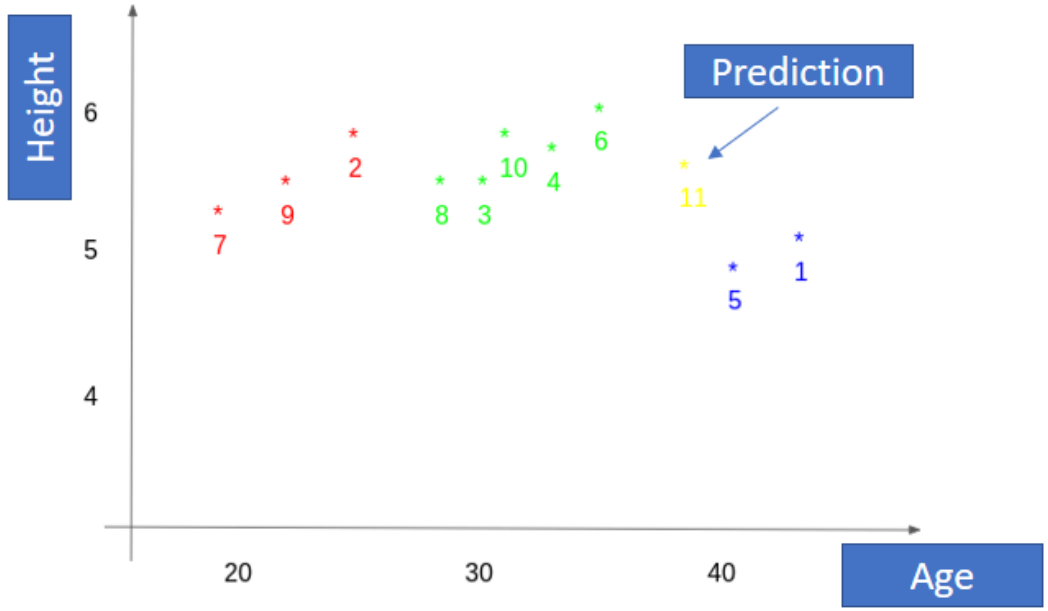
Image: Google.com

Let's Calculate for Regression!

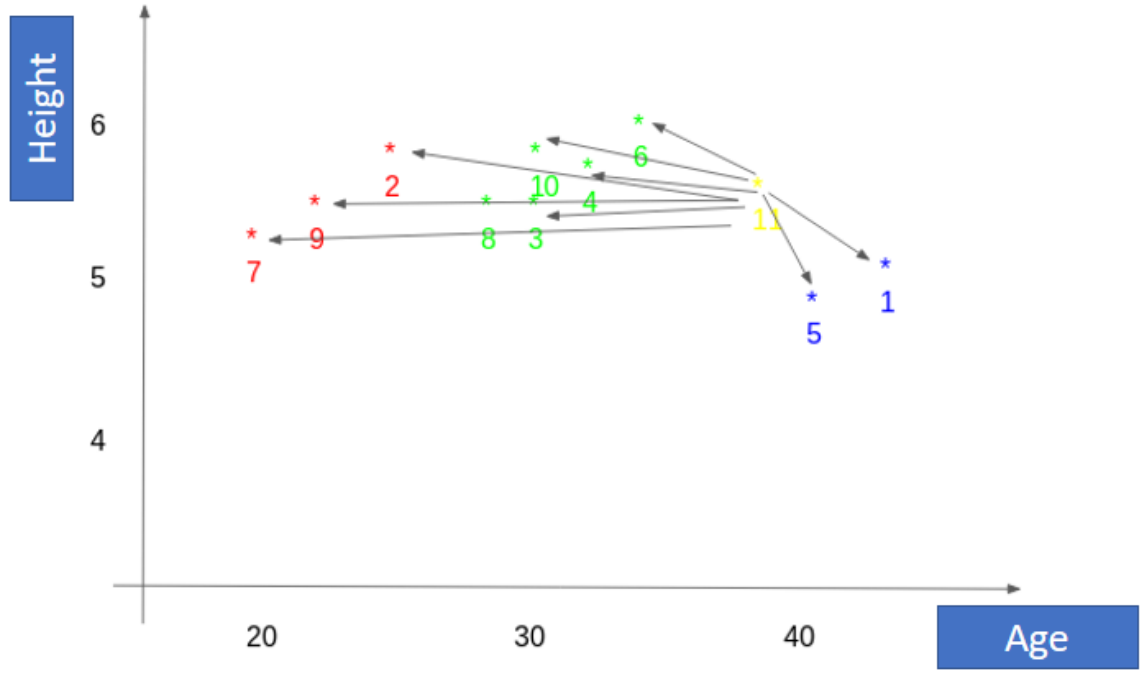
K-NN Algorithm: Regression

A	B	C	D
id	age	height	weight
1	45	5	77
2	26	5.11	47
3	30	5.6	55
4	34	5.9	59
5	40	4.8	72
6	36	5.8	60
7	19	5.3	40
8	28	5.8	60
9	23	5.5	45
10	32	5.6	58
11	38	5.5	?

K-NN Algorithm: Regression



K-NN Algorithm: Regression



K-NN Algorithm: Regression

	A	B	C	D
1	id	age	height	weight
2	1	45	5	77
3	2	26	5.11	47
4	3	30	5.6	55
5	4	34	5.9	59
6	5	40	4.8	72
7	6	36	5.8	60
8	7	19	5.3	40
9	8	28	5.8	60
10	9	23	5.5	45
11	10	32	5.6	58
12		38	5.5	?

Distance,

$$d(p,q)^2 = (q_1 - p_1)^2 + (q_2 - p_2)^2$$

	A	B	C	E
1	id	age	height	distance
2	1	45	5	a
3	2	26	5.11	b
4	3	30	5.6	c
5	4	34	5.9	d
6	5	40	4.8	e
7	6	36	5.8	f
8	7	19	5.3	g
9	8	28	5.8	h
10	9	23	5.5	i
11	10	32	5.6	j
12	11	38	5.5	
13				

K-NN Algorithm: Regression

	A	B	C	D
1	id	age	height	weight
2	1	45	5	77
3	2	26	5.11	47
4	3	30	5.6	55
5	4	34	5.9	59
6	5	40	4.8	72
7	6	36	5.8	60
8	7	19	5.3	40
9	8	28	5.8	60
10	9	23	5.5	45
11	10	32	5.6	58
12		38	5.5	?

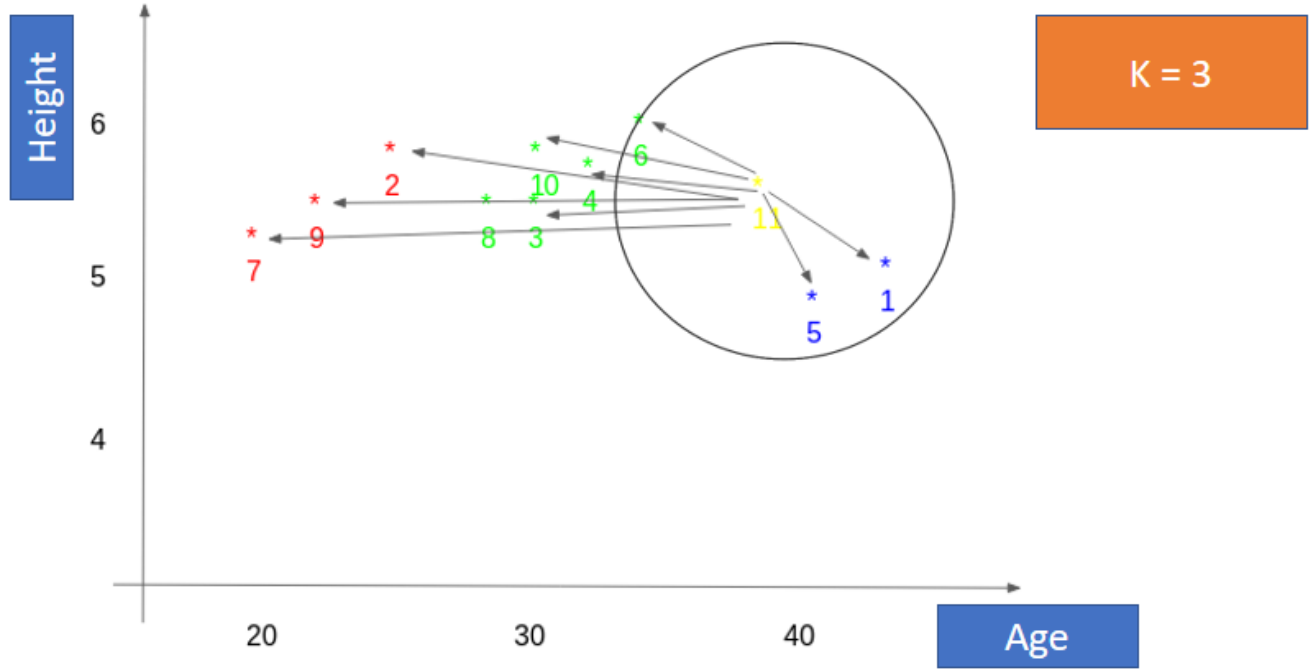
Distance,

$$d(p,q)^2 = (q_1 - p_1)^2 + (q_2 - p_2)^2$$

Sequence:

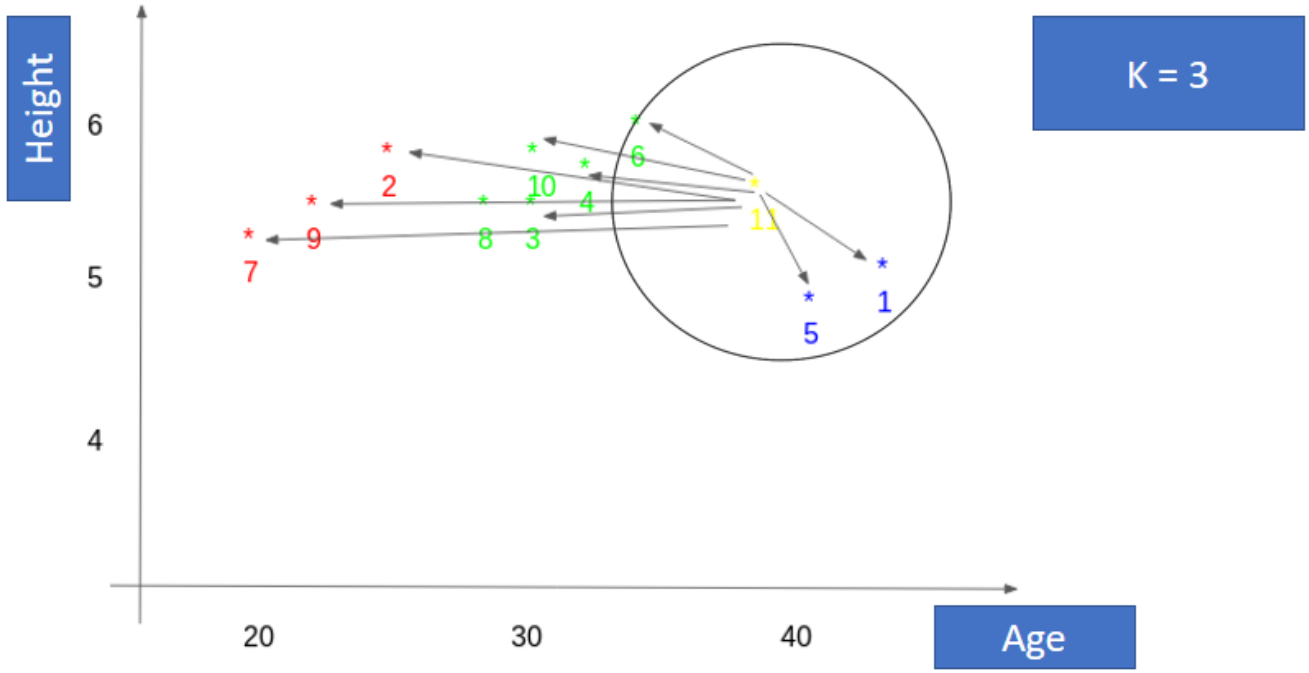
a > e > f > d > j > c > h > b > i > g

K-NN Algorithm: Regression



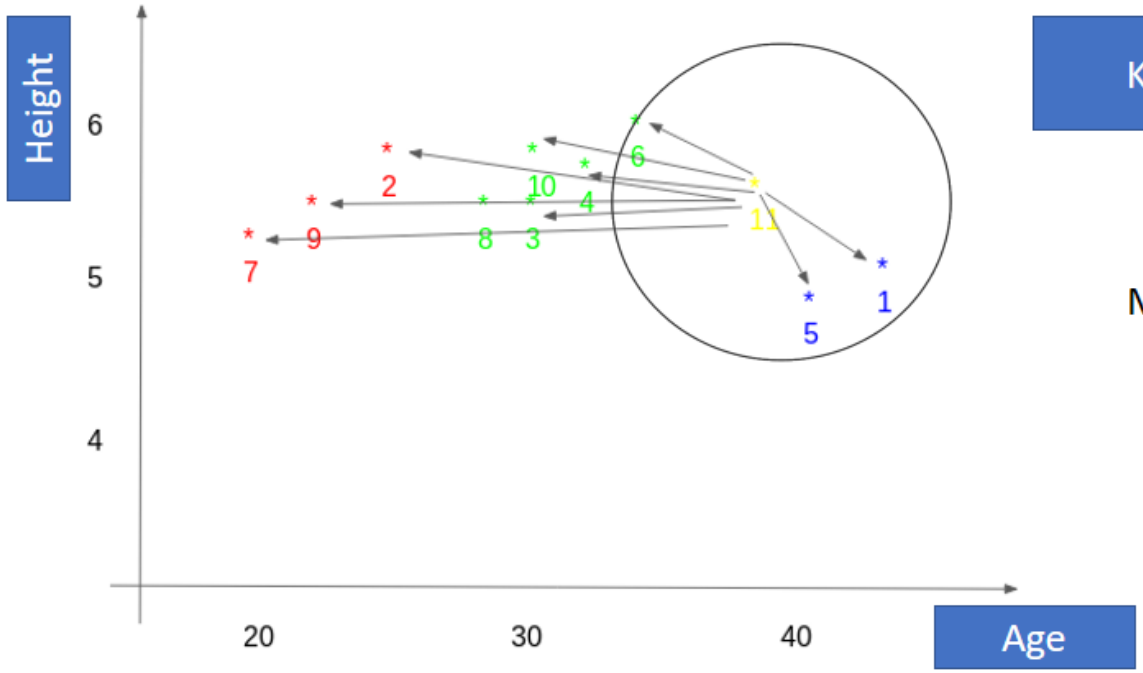
K-NN Algorithm: Regression

A	B	C	D
id	age	height	weight
1	45	5	77
2	26	5.11	47
3	30	5.6	55
4	34	5.9	59
5	40	4.8	72
6	36	5.8	60
7	19	5.3	40
8	28	5.8	60
9	23	5.5	45
10	32	5.6	58
	38	5.5	?



K-NN Algorithm: Regression

A	B	C	D
id	age	height	weight
1	45	5	77
2	26	5.11	47
3	30	5.6	55
4	34	5.9	59
5	40	4.8	72
6	36	5.8	60
7	19	5.3	40
8	28	5.8	60
9	23	5.5	45
10	32	5.6	58
	38	5.5	?



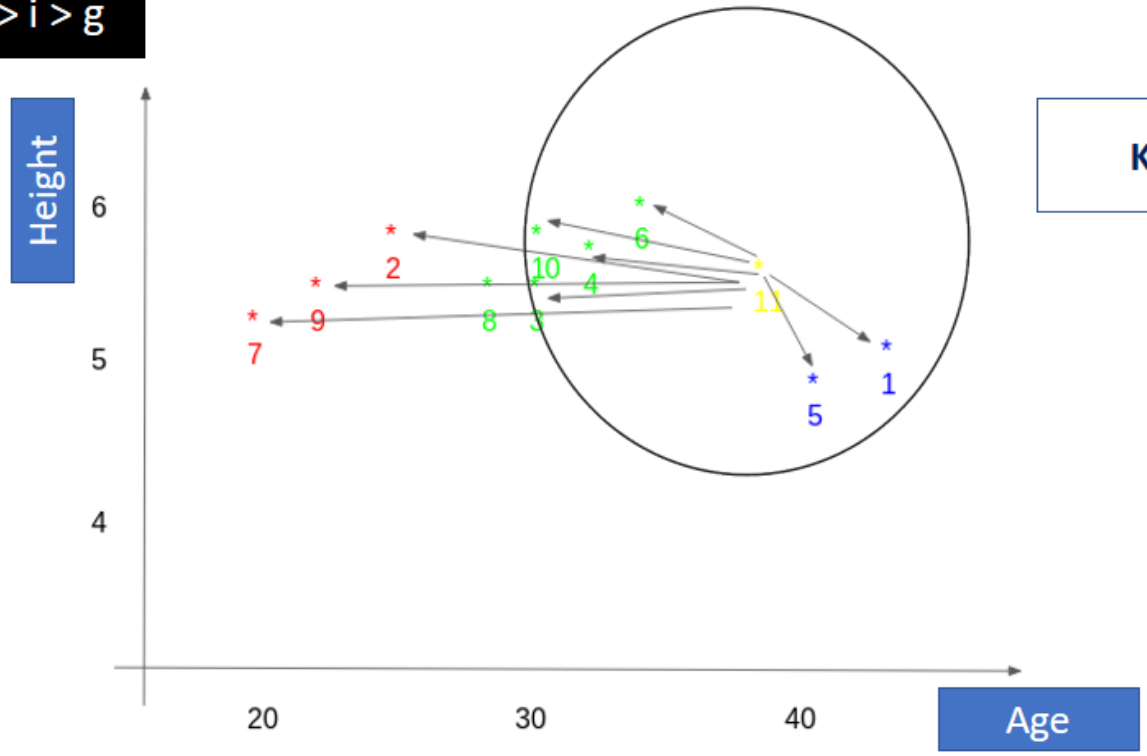
K = 3

Mean: (77+72+60)/3
= 69.6 kg

K-NN Algorithm: Regression

Sequence:
a > e > f > d > j > c > h > b > i > g

	A	B	C	D	E
1	id	age	height	weight	distance
2	1	45	5	77	a
3	2	26	5.11	47	b
4	3	30	5.6	55	c
5	4	34	5.9	59	d
6	5	40	4.8	72	e
7	6	36	5.8	60	f
8	7	19	5.3	40	g
9	8	28	5.8	60	h
10	9	23	5.5	45	i
11	10	32	5.6	58	j
12	11	38	5.5	?	
13					

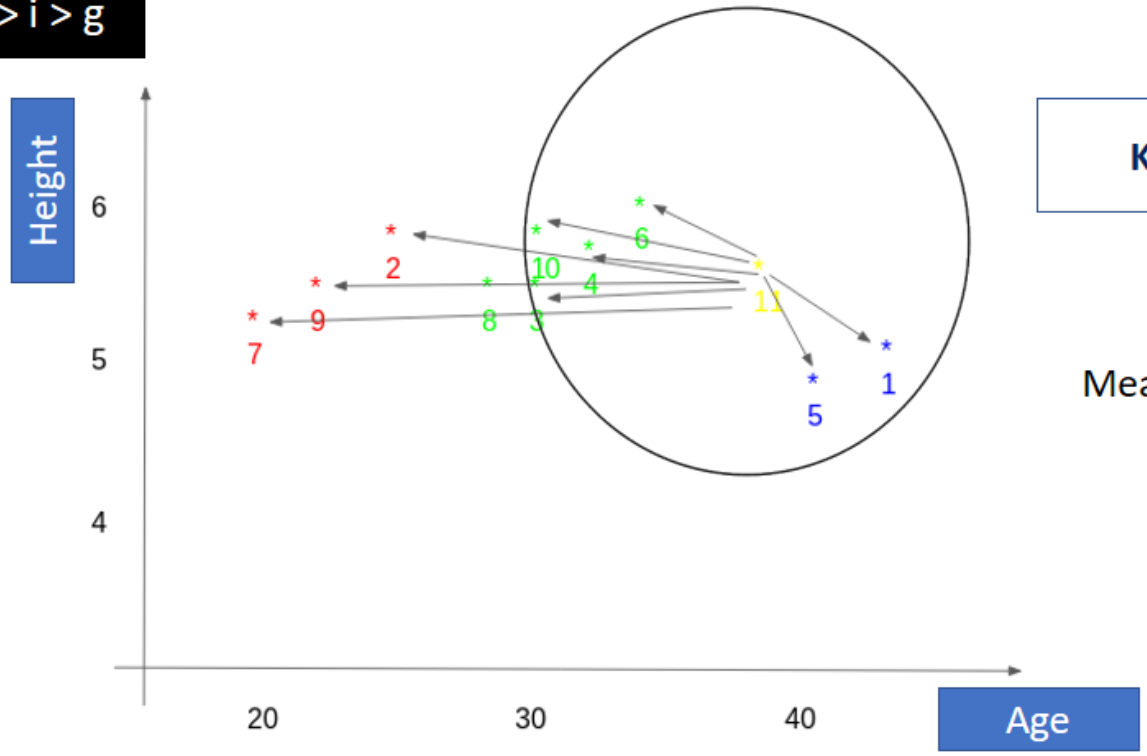


K = 5

K-NN Algorithm: Regression

Sequence:
a > e > f > d > j > c > h > b > i > g

	A	B	C	D	E
1	id	age	height	weight	distance
2	1	45	5	77	a
3	2	26	5.11	47	b
4	3	30	5.6	55	c
5	4	34	5.9	59	d
6	5	40	4.8	72	e
7	6	36	5.8	60	f
8	7	19	5.3	40	g
9	8	28	5.8	60	h
10	9	23	5.5	45	i
11	10	32	5.6	58	j
12	11	38	5.5	?	
13					



K = 5

Mean: (77+72+60+59+58)/5
= 65.2 kg

Let's Do Self Assessment!