How does k-NN work?

Consider the dataset having two variables: **height(cm)** and **weight(kg)** and each datapoint is classified as **Normal** or **Underweight**

Weight(x2)	Height(y2)	Class
51	167	Underweight
62	182	Normal
69	176	Normal
64	173	Normal
65	172	Normal
56	174	Underweight
58	169	Normal
57	173	Normal
55	170	Normal

How does k-NN work?

Assuming we do not know how to calculate the **body mass index (BMI)**, we need to classify this unseen data as **Normal** or **Underweight**

57 kg	170 cm	?

How does k-NN work?

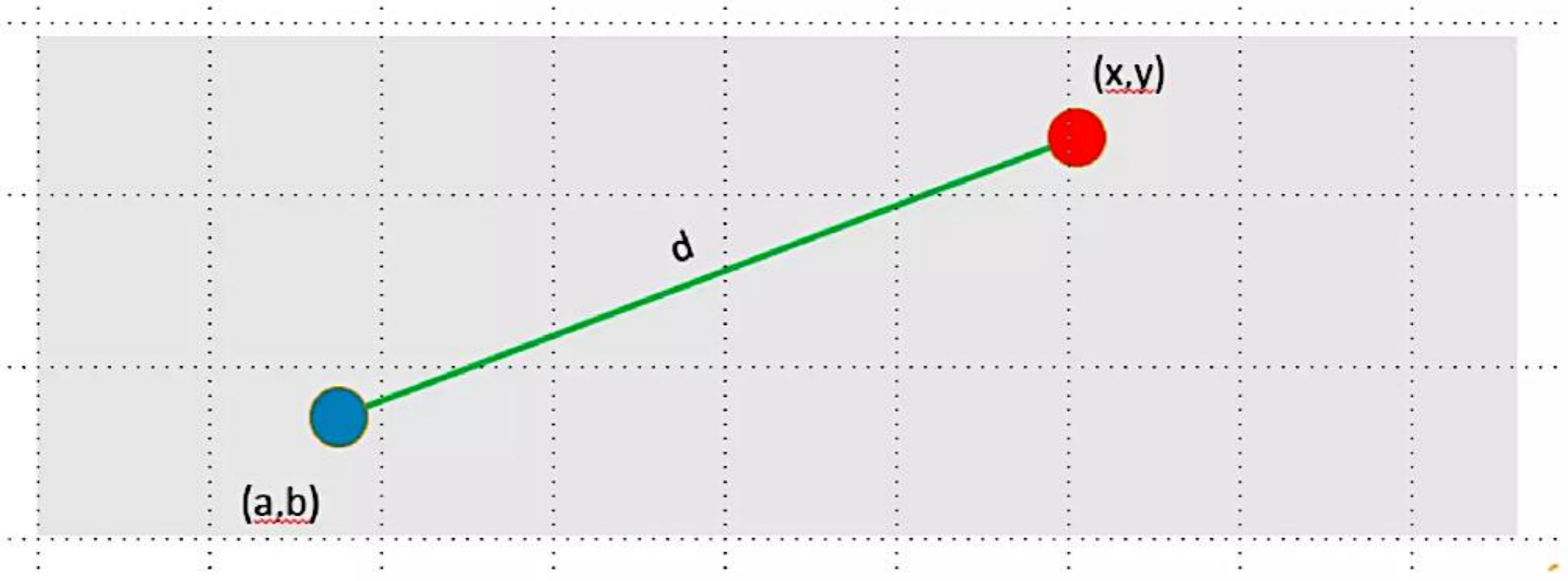
KNN finds the nearest neighbor by calculating the **Euclidean distance**

57 kg	170 cm	?

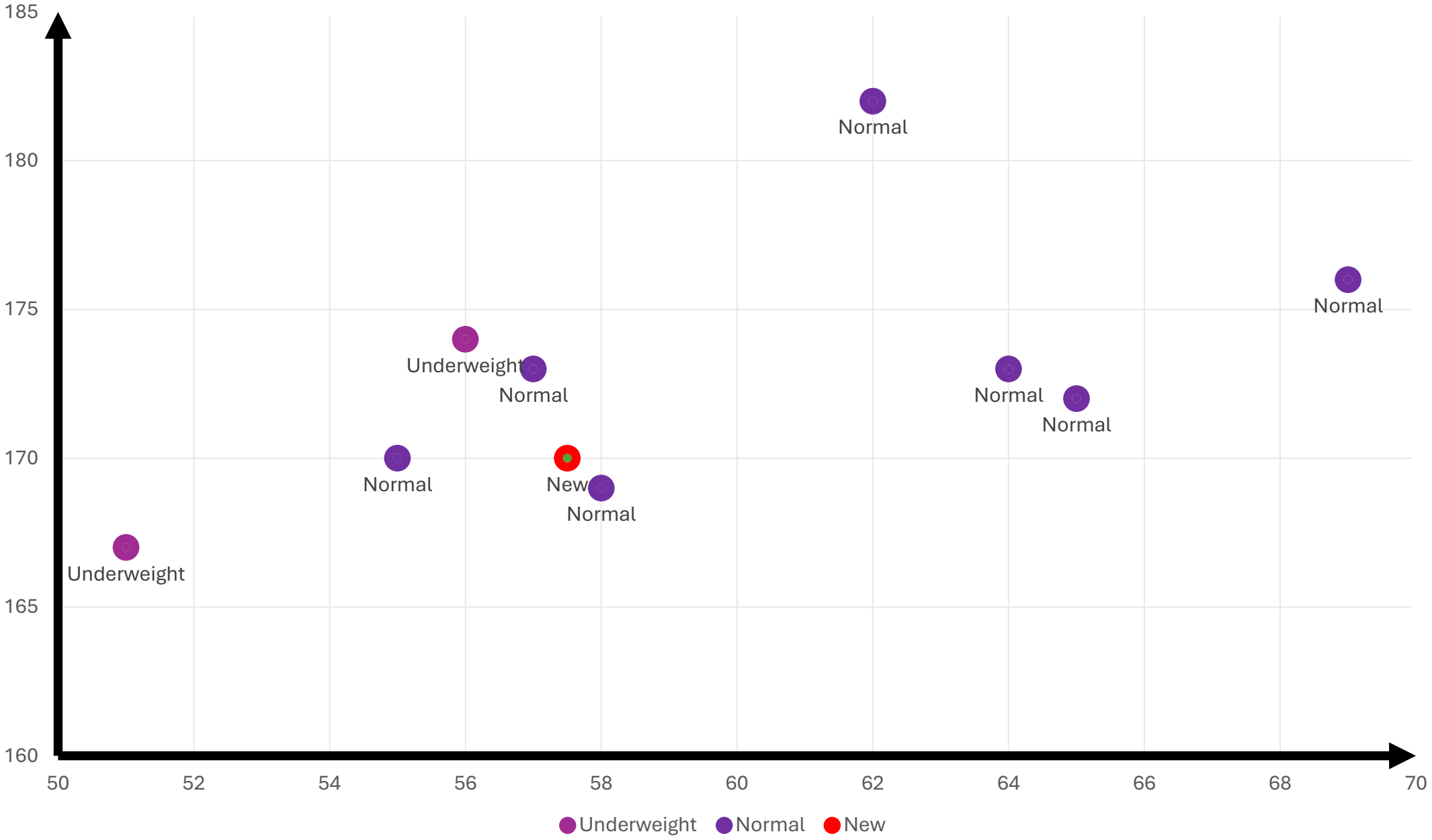
How does k-NN work?

k-NN uses **Euclidean distance formula**, to calculate the distance between two data points

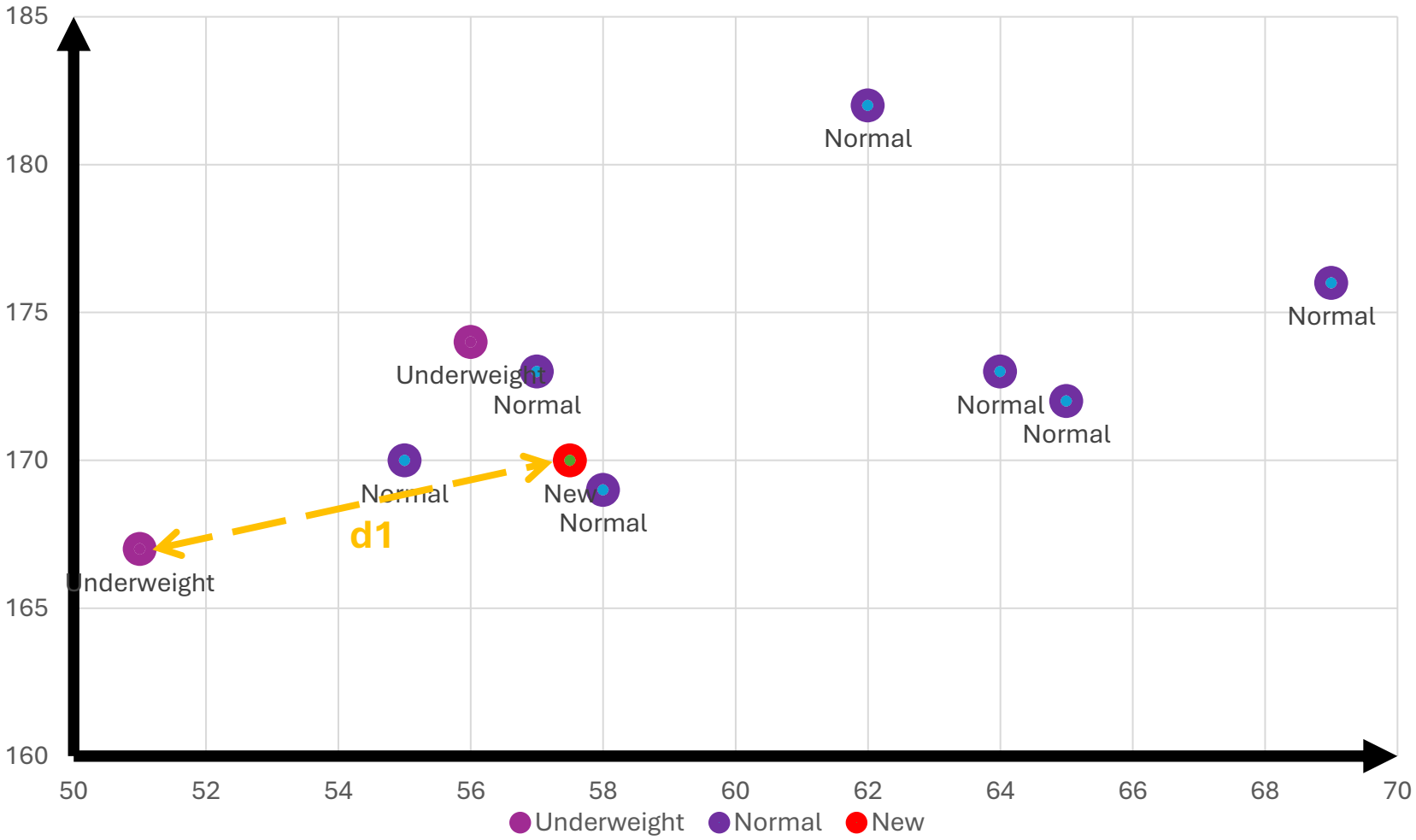
$$d = \sqrt{(x-a)^2+(y-b)^2}$$



$$d = \sqrt{(x-a)^2+(y-b)^2}$$

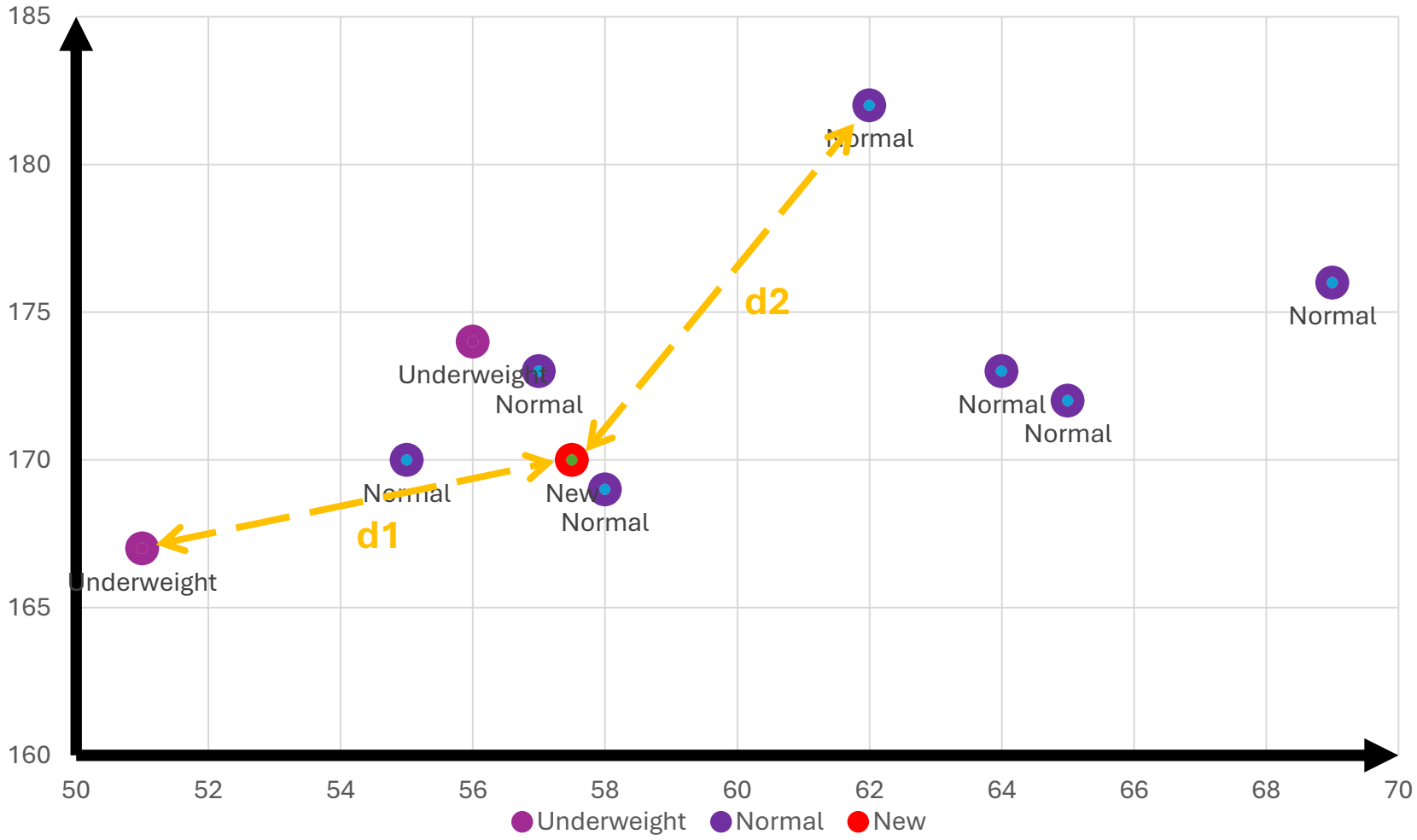

185
180
175
170
165
160
50
52
54
56
58
60
62
64
66
68
70
Normal
Normal
Underweight
Normal
Normal
Normal
Normal
New
Normal
Underweight
Underweight
Normal
New

CCMACLRL



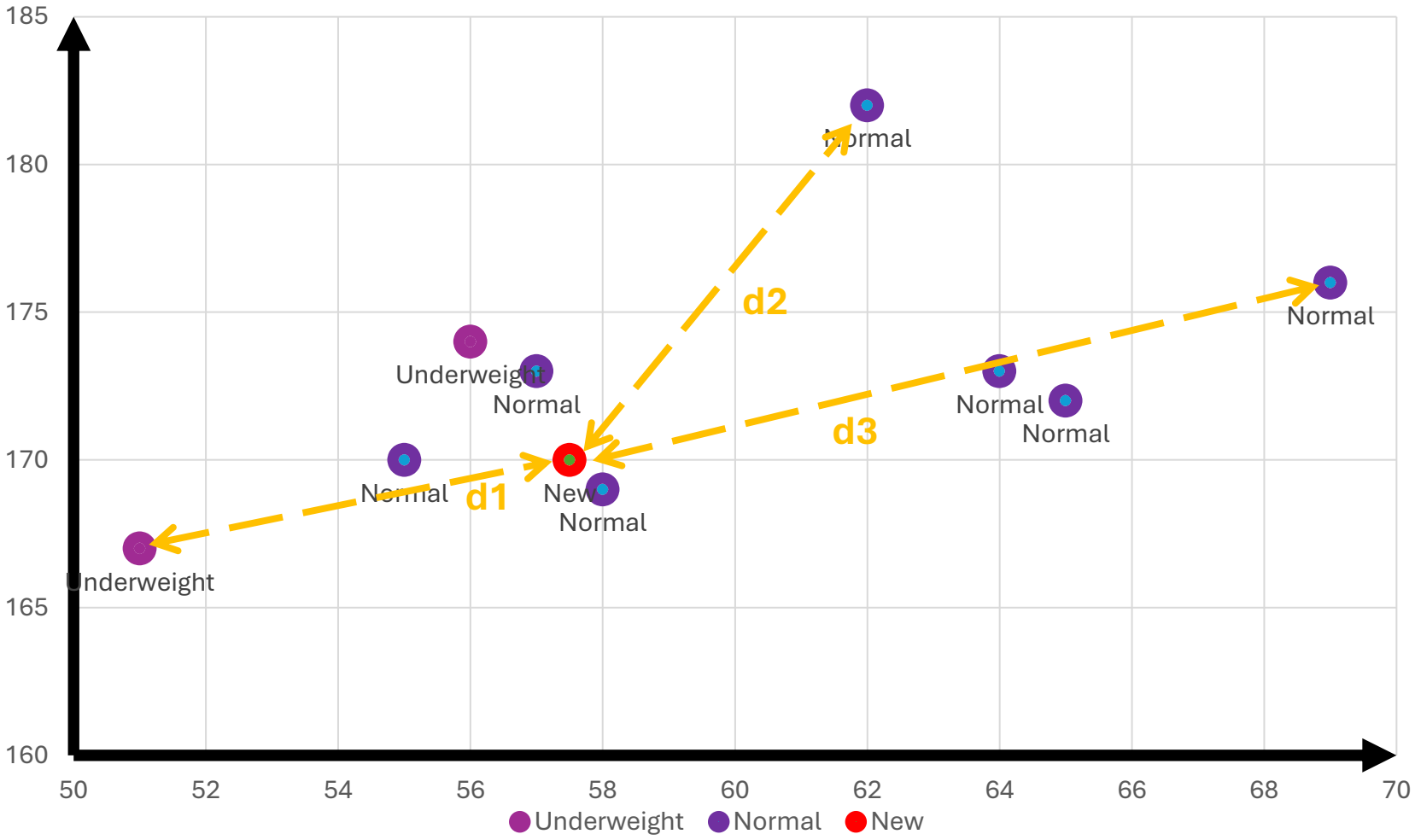
$$d1 = \sqrt{(170-167)^2 - (57-51)^2}$$

$$d1 = 6.7$$



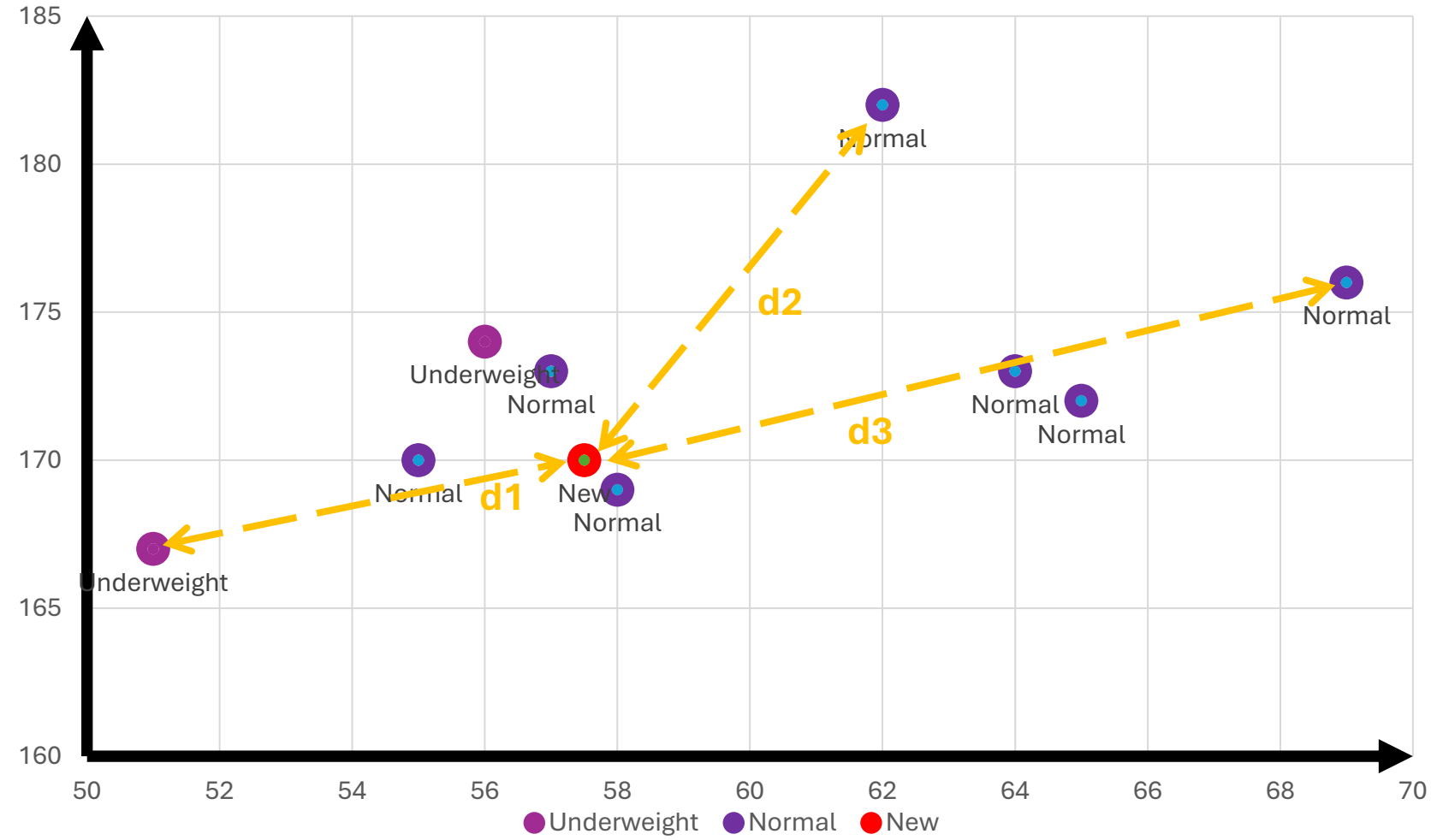
$$d2 = \sqrt{(170-182)^2 - (57-62)^2}$$

$$d2 = 13$$



$$d3 = \sqrt{(170-176)^2 - (57-69)^2}$$

$$d3 = 13.4$$



$$d1 = \sqrt{(170-167)^2 - (57-51)^2}$$
$$d1 = 6.7$$

$$d2 = \sqrt{(170-182)^2 - (57-62)^2}$$
$$d2 = 13$$

$$d3 = \sqrt{(170-176)^2 - (57-69)^2}$$
$$d3 = 13.4$$

Similarly, we calculate the **Euclidean distance** of the **unknown data** point from all points in the dataset

How does k-NN work?

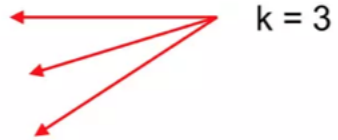
After we calculate the **Euclidean distance** of the unknown datapoint from all the data points, we get...

Weight(x2)	Height(y2)	Class	Euclidean Distance
51	167	Underweight	6.7
62	182	Normal	13
69	176	Normal	13.4
64	173	Normal	7.6
65	172	Normal	8.2
56	174	Underweight	4.1
58	169	Normal	1.4
57	173	Normal	3
55	170	Normal	2

How does k-NN work?

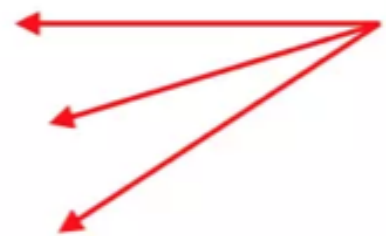
Let us calculate the nearest neighbor at **k=3**

Weight(x2)	Height(y2)	Class	Euclidean Distance
51	167	Underweight	6.7
62	182	Normal	13
69	176	Normal	13.4
64	173	Normal	7.6
65	172	Normal	8.2
56	174	Underweight	4.1
58	169	Normal	1.4
57	173	Normal	3
55	170	Normal	2



57 kg	170 cm	?

Weight(x2)	Height(y2)	Class	Euclidean Distance
51	167	Underweight	6.7
62	182	Normal	13
69	176	Normal	13.4
64	173	Normal	7.6
65	172	Normal	8.2
56	174	Underweight	4.1
58	169	Normal	1.4
57	173	Normal	3
55	170	Normal	2



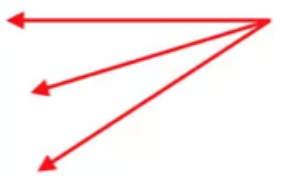
k = 3

57 kg	170 cm	?

The majority of its nearest neighbors are **Normal**.

Hence, having a weight of **57 kg** and a height of **170 cm** is classified as **Normal**

Class	Euclidean Distance
Underweight	6.7
Normal	13
Normal	13.4
Normal	7.6
Normal	8.2
Underweight	4.1
Normal	1.4
Normal	3
Normal	2



k = 3

57 kg	170 cm	**Normal**