

1. The table below lists a dataset that was used to create a nearest neighbour model that predicts whether it will be a good day to go surfing.

ID	WAVE SIZE (FT)	WAVE PERIOD (SECS)	WIND SPEED (MPH)	GOOD SURF
1	6	15	5	yes
2	1	6	9	no
3	7	10	4	yes
4	7	12	3	yes
5	2	2	10	no
6	10	2	20	no

Assuming that the model uses Euclidean distance to find the nearest neighbour, what prediction will the model return for each of the following query instances.

ID	WAVE SIZE (FT)	WAVE PERIOD (SECS)	WIND SPEED (MPH)	GOOD SURF
Q1	8	15	2	?
Q2	8	2	18	?
Q3	6	11	4	?

The table below lists the training instances along with the distances between each training instance and each query. The distance between each query instance and its nearest training instance is highlighted in bold.

ID	WAVE SIZE (FT)	WAVE PERIOD (SECS)	WIND SPEED (MPH)	GOOD SURF	Euc. Dist. to Q1	Euc. Dist. to Q2	Euc. Dist. to Q3
1	6	15	5	yes	3.61	18.49	4.12
2	1	6	9	no	13.38	12.08	8.66
3	7	10	4	yes	5.48	16.16	1.41
4	7	12	3	yes	3.32	18.06	1.73
5	2	2	10	no	16.40	10.00	11.53
6	10	2	20	no	22.29	2.83	18.79

From this table we can see that

- The nearest neighbour to Q1 is training instance **d₄** which is 3.32 units away from Q1. This training instance has a target level of GOOD SURF=yes. So the model will predict GOOD SURF=yes for Q1.
- The nearest neighbour to Q2 is training instance **d₆** which is 2.83 units away from Q2. This training instance has a target level of GOOD SURF=no. So the model will predict GOOD SURF=no for Q2.
- The nearest neighbour to Q3 is training instance **d₃** which is 1.41 units away from Q3. This training instance has a target level of GOOD SURF=yes. So the model will predict GOOD SURF=yes for Q3.