

Assignment 1

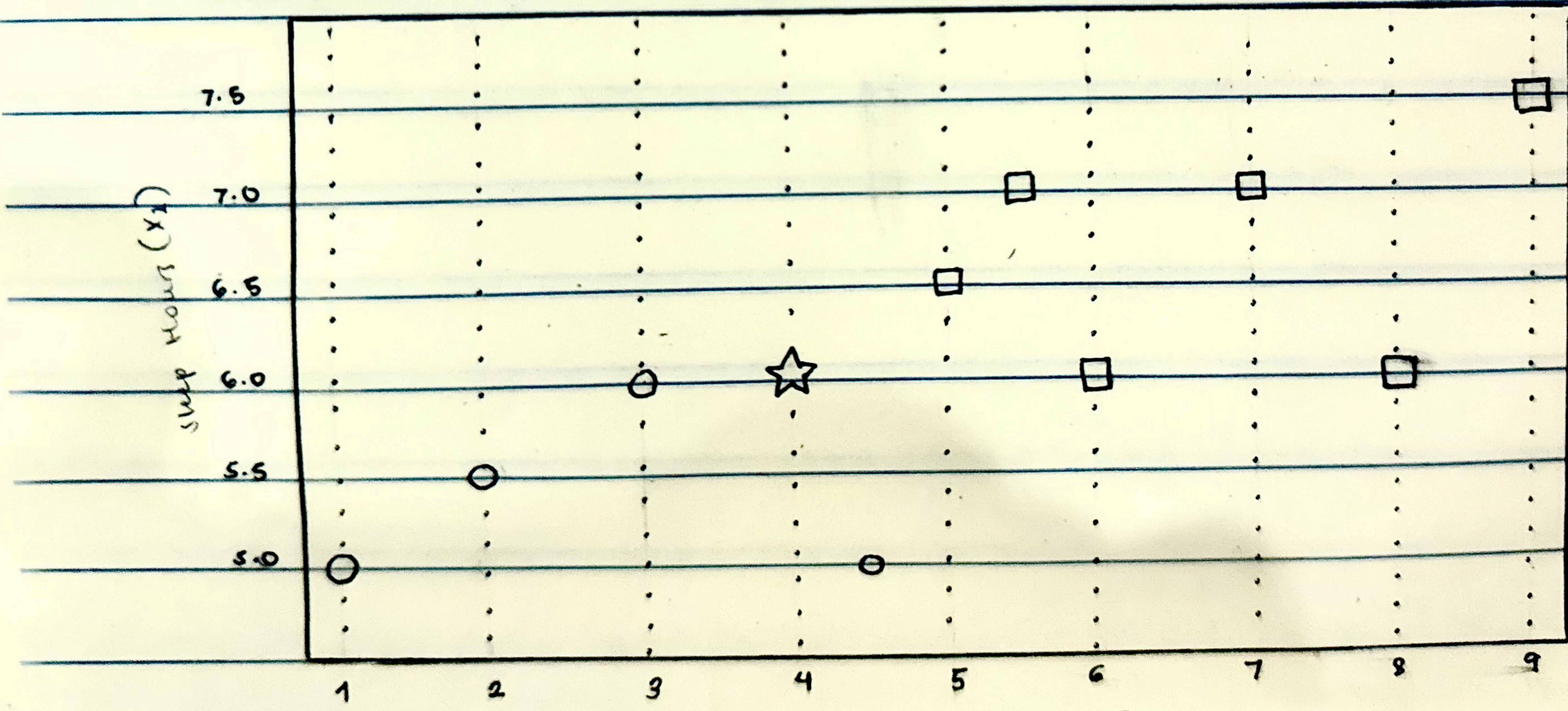
using k-Nearest Neighbors , predict if a student will pass(1) or fail(0) based on :

X_1 : Hours Studied

X_2 : Sleep Hours

1. Plot all data points using a scatter plot : Failed - Circle ; Passed - Square ; Unknown - Star

KNN classification



Hours Studied (X_1)

② Compute the distance

$$d = \sqrt{(x_1 - 4)^2 + (x_2 - 6)^2}$$

1. Student 1 (1, 5) :

$$= \sqrt{(1-4)^2 + (5-6)^2}$$

$$= \sqrt{9+1}$$

$$= \sqrt{10}$$

$$= 3.162$$

2. student 2 (2, 5.5) :

$$= \sqrt{(2-4)^2 + (5.5-6)^2}$$

$$= \sqrt{4 + 0.25}$$

$$= \sqrt{4.25}$$

$$= 2.062$$

3. student 3 (3, 6) :

$$= \sqrt{(3-4)^2 + (6-6)^2}$$

$$= \sqrt{1+0}$$

$$= \sqrt{1}$$

$$= 1.000$$

4. student 4 (4.5, 5) :

$$= \sqrt{(4.5-4)^2 + (5-6)^2}$$

$$= \sqrt{0.25+1}$$

$$= \sqrt{1.25}$$

$$= 1.118$$

5. student 5 (5, 6.5) :

$$= \sqrt{(5-4)^2 + (6.5-6)^2}$$

$$= \sqrt{1+0.25}$$

$$= \sqrt{1.25}$$

$$= 1.118$$

6. student 6 (5.5, 7) :

$$= \sqrt{(5.5-4)^2 + (7-6)^2}$$

$$= \sqrt{2.25+1}$$

$$= \sqrt{3.25}$$

$$= 1.803$$

7. student 7 (6, 6) :

$$= \sqrt{(6-4)^2 + (6-6)^2}$$

$$= \sqrt{4+0}$$

$$= \sqrt{4}$$

$$= 2.000$$

8. student 8 (7, 7) :

$$= \sqrt{(7-4)^2 + (7-6)^2}$$

$$= \sqrt{9+1}$$

$$= \sqrt{10}$$

$$= 3.162$$

9. student 9 (8, 6)

$$= \sqrt{(8-4)^2 + (6-6)^2}$$

$$= \sqrt{16+0}$$

$$= \sqrt{16}$$

$$= 4.000$$

10. student 10 (9, 7.5)

$$= \sqrt{(9-4)^2 + (7.5-6)^2}$$

$$= \sqrt{25+2.25}$$

$$= \sqrt{27.25}$$

$$= 5.220$$

② Fill the table

Student	Hours studied (x_1)	Sleep Hours (x_2)	Pass / Fail (y)	Euclidean Distance
1	1.0	5.0	0	3.162
2	2.0	5.5	0	2.062
3	3.0	6.0	0	1.000
4	4.5	5.0	0	1.118
5	5.0	6.5	1	1.118
6	5.5	7.0	1	1.803
7	6.0	6.0	1	2.000
8	7.0	7.0	1	3.162
9	8.0	6.0	1	4.000
10	9.0	7.5	1	5.220

2. Find the 3 Nearest Neighbors

Student	Hours Studied (x_1)	Sleep Hours (x_2)	Pass / Fail (y)	Euclidean Distance
1	1.0	5.0	0	3.162
2	2.0	5.5	0	2.062
3	3.0	6.0	0	1.000
4	4.5	5.0	0	1.118
5	5.0	6.5	1	1.118
6	5.5	7.0	1	1.803
7	6.0	6.0	1	2.000
8	7.0	7.0	1	3.162
9	8.0	6.0	1	4.000
10	9.0	7.5	1	5.220

③ Majority Vote

Count how many are Pass(1) and how many are Fail(0)

Student 3 \rightarrow Fail(0) \rightarrow Distance = 1.000 Fail = 2

Student 4 \rightarrow Fail(0) \rightarrow Distance = 1.118 Pass = 1

Student 5 \rightarrow Pass(1) \rightarrow Distance = 1.118 Majority = Fail

Prediction: A new student who studied 4 hours and slept 6 hours
will Fail the exam.

④ Discussion Questions

1. What was your final prediction?

\hookrightarrow My final prediction is that if the new student studied for four(4) hours and slept for six(6) hours they are more likely to fail the exam.

2. How would the prediction change if we used $k=5$ instead of $k=3$?

\hookrightarrow The prediction would change as the data suggests otherwise on $k=3$, Based on the table, from students three(3) to seven(7), the majority students passed the exam. therefore, the prediction would be the new student is more likely to pass the exam.