NATIONAL UNIVERSITY - MANILA
COLLEGE OF COMPUTING AND INFORMATION TECHNOLOGY

Introduction to Machine Learning

K-NEAREST NEIGHBOURS

(PREDICTING PASS OR FAIL)

ASSIGNMENT NO. 1

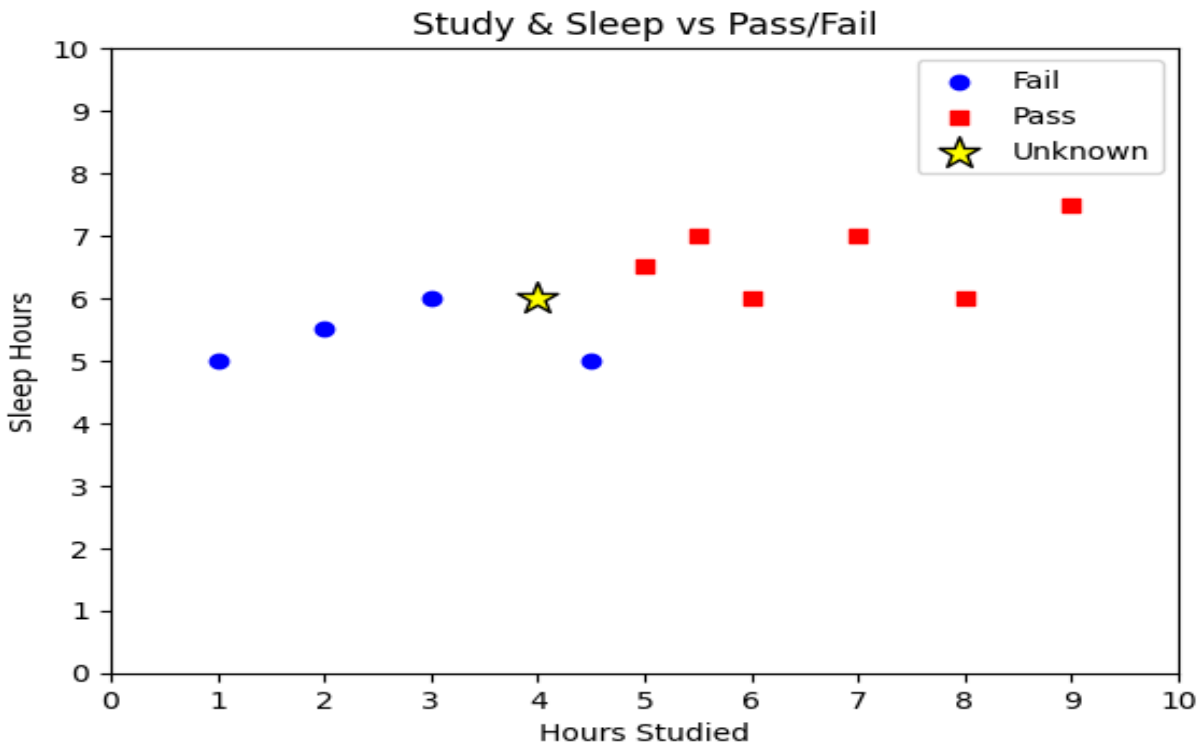
Prepared by:

Rae S. Paulos

TRAINING DATA (10 STUDENTS)

Student	Hours Studied (X1)	Sleep Hours (X2)	Pass/Fail (Y)
1	1.0	5.0	0 (Fail)
2	2.0	5.5	0 (Fail)
3	3.0	6.0	0 (Fail)
4	4.5	5.5	0 (Fail)
5	5.0	6.5	1 (Pass)
6	5.5	7.0	1 (Pass)
7	6.0	6.0	1 (Pass)
8	7.0	7.0	1 (Pass)
9	8.0	6.0	1 (Pass)
10	9.0	7.5	1 (Pass)

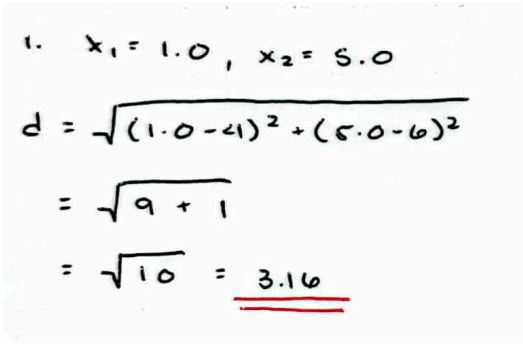
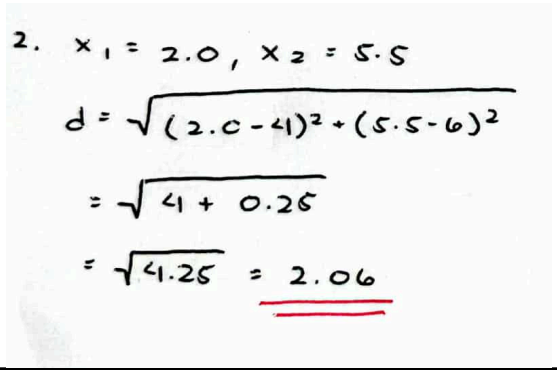
1. **Plot all data using a scatter plot (10 points)**

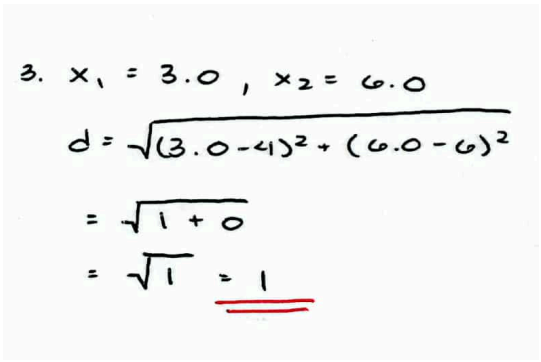
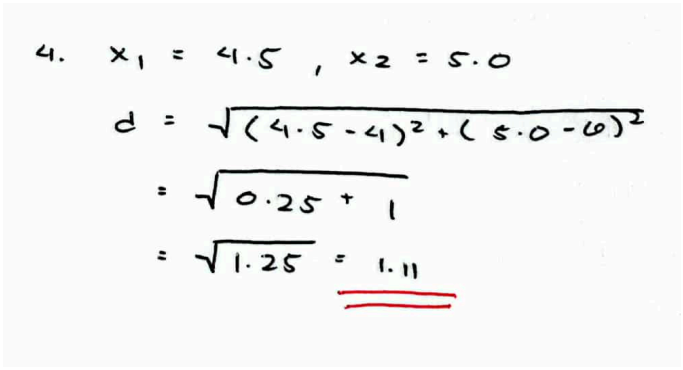
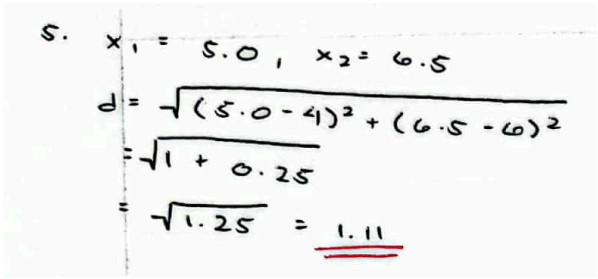
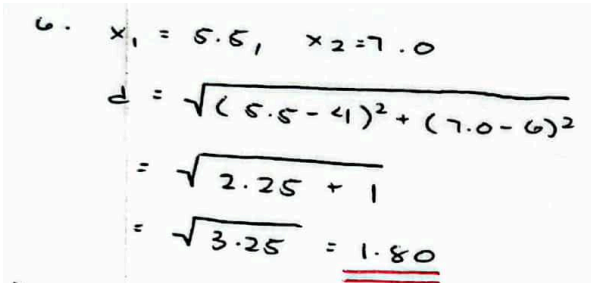
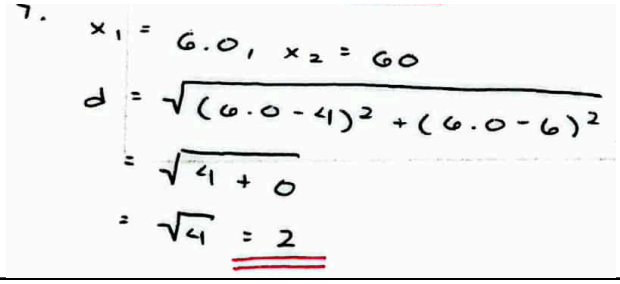


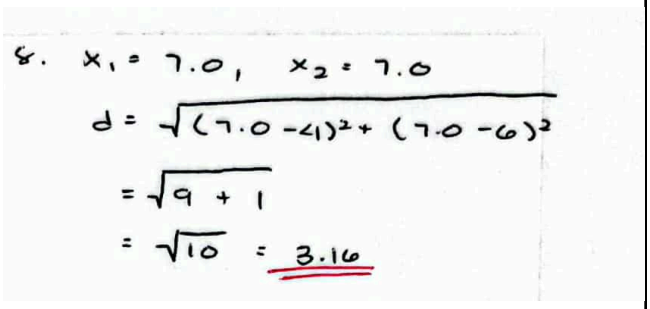
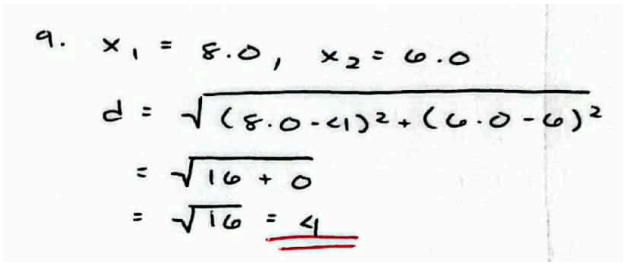
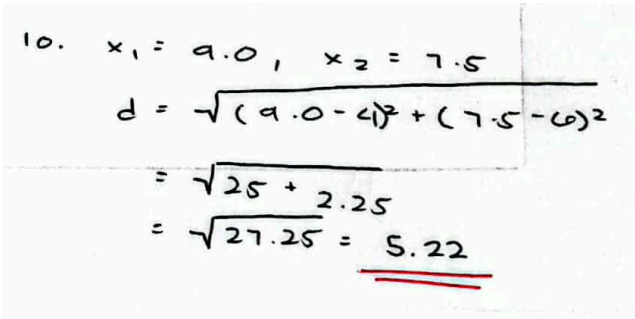
2. Compute Distances (10 points)

Student	Hours Studied (X1)	Sleep Hours (X2)	Pass/Fail (Y)	Euclidean Distance
1	1.0	5.0	0 (Fail)	**3.16**
2	2.0	5.5	0 (Fail)	**2.06**
3	3.0	6.0	0 (Fail)	**1.0**
4	4.5	5.5	0 (Fail)	**1.11**
5	5.0	6.5	1 (Pass)	**1.11**
6	5.5	7.0	1 (Pass)	**1.80**
7	6.0	6.0	1 (Pass)	**2.0**
8	7.0	7.0	1 (Pass)	**3.16**
9	8.0	6.0	1 (Pass)	**4.0**
10	9.0	7.5	1 (Pass)	**5.22**

Solution:

#	Solution	Answer
1		**3.16**
2		**2.06**

3		**1.0**
4		**1.11**
5		**1.11**
6		**1.80**
7		**2.0**

8		**3.16**
9		**4.0**
10		**5.22**

3. Find the 3 Nearest Neighbors (3 points)

Highlight the rows with the 3 smallest distance

Student	Hours Studied (X1)	Sleep Hours (X2)	Pass/Fail (Y)	Euclidean Distance
1	1.0	5.0	0 (Fail)	**3.16**
2	2.0	5.5	0 (Fail)	**2.06**
3	3.0	6.0	0 (Fail)	**1.0**
4	4.5	5.5	0 (Fail)	**1.11**
5	5.0	6.5	1 (Pass)	**1.11**
6	5.5	7.0	1 (Pass)	**1.80**

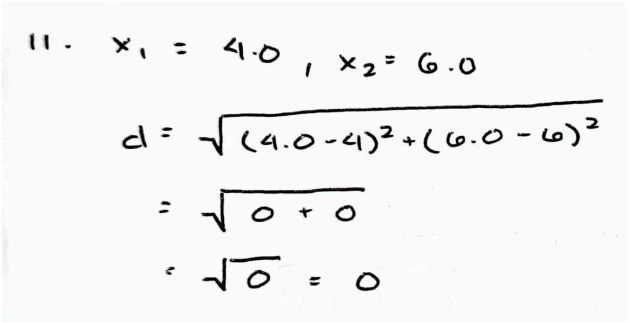
7	6.0	6.0	1 (Pass)	**2.0**
8	7.0	7.0	1 (Pass)	**3.16**
9	8.0	6.0	1 (Pass)	**4.0**
10	9.0	7.5	1 (Pass)	**5.22**

4. Majority Vote (2 points)

Count how many passed and how many failed. Predict the outcome for the new student who studied 4 hours and slept 6 hours.

Category	Count
Pass	6 students
Fail	4 students

Euclidean Distance

#	Solution	Answer
11		**0**

Prediction:

Student	Hours Studied (X1)	Sleep Hours (X2)	Pass/Fail (Y)	Euclidean Distance
11	4.0	6.0	**0 (Fail)**	**0**

5. Discussion Questions (5 points)

Question	Answer
What was your final prediction?	The final prediction for a student who

	slept for 6 hours and studied for 4 hours **failed**.
How would the prediction change if we used k = 5 instead of k = 3?	With k = 3, the student is predicted to fail, but **with k = 5, the student is predicted to pass**. This change occurs because **increasing k shifts the decision from relying on just a few nearby points to considering a wider group of neighbors.** In the smaller neighborhood, the majority outcome was failure, but when two more neighbors were added, the balance tipped toward passing. This illustrates how the choice of k directly influences the classification result.